DOCUMENT RESUME

ED 394 658 PS 024 022

AUTHOR	Thomas, Nancy G., Ed.
TITLE	Social Policy Report, 1993.
INSTITUTION	Society for Research in Child Development.
PUB DATE	93
NOTE	101p.; For 1994 reports, see PS 024 095.
AVAILABLE FROM	SRCD Executive Office, University of Michigan, 300 North Ingalls, 10th Floor, Ann Arbor, MI 48109-0406 (Single issues, $3.50 each; Subscription to nonmembers of SRCD, $12.50).
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	Social Policy Report; v7 n1-4 Spr-Win 1993
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	Bilingualism; *Children; Court Litigation; Cultural Pluralism; Ethics; Foreign Countries; Learning Disabilities; Program Evaluation; *Public Policy; *Research Administration; Research Problems; Sexual Abuse; *Special Education; Theory Practice Relationship
IDENTIFIERS	Canada; Informed Consent; *Project Head Start; *Witnesses

ABSTRACT

Each of the four issues of this newsletter published in 1993 consists of one article dealing with a particular policy debate. Number 1, "Canadian Special Education Policies: Children with Learning Disabilities in a Bilingual and Multicultural Society" (Linda S. Siegel and Judith Wiener), discusses social and cultural factors affecting the education of children with learning disabilities, including policy issues, and bilingualism and multiculturalism. This issue contains 34 references. Number 2, "Using Research and Theory To Justify and Inform Head Start Expansion" (Edward Zigler and Sally J. Styfco), examines Head Start's programs and goals, and addresses issues surrounding the evaluation of Head Start, policy directions (including issues of quantity and quality), and planning for a better, bigger Head Start program. This issue contains 87 references. Number 3, "Child Witnesses: Translating Research into Policy" (Stephen J. Ceci and Maggie Bruck), begins with a discussion of the McMartin preschool (Manhattan Beach, California) child sexual abuse case, and moves on to a discussion of the prevalence of abuse and court involvement. Research on children's suggestibility, including past and present trends, are discussed as are other issues, including policy implications for expert witnesses, the relationship of research to clinical practice, and professional organizations as ethical gatekeepers. This issue contains 97 references. Number 4, "Integrating Science and Ethics in Research with High-Risk Children and Youth (Celia B. Fisher), deals with federal and professional guidelines for research with children and youth, and such issues as informed consent, problems identified during the course of research, generalization of research results to individuals and diverse populations, and recommendations for future policy. This issue contains 102 references. (DR)

SOCIAL POLICY REPORT
Society For Research In Child Development

Volume VII, Number 1
Spring 1993

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 394 658

Canadian Special Education Policies: Children with Learning Disabilities in a Bilingual and Multicultural Society

Linda S. Siegel
Judith Wiener

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Nancy G. Thomas
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Introduction

Canada is a vast and diverse country with a relatively small population of approximately 27 million. Although on a superficial level it may seem similar to the United States in economic structure, language, and culture, we will argue in this report that there are some important differences. We will discuss how societal and cultural factors affect the education of children, particularly those with learning disabilities. Of special interest will be the role of bilingualism and multiculturalism in Canada's approach to the education of its children, and we will focus on areas that distinguish Canada from the U.S.

Present-day Canada has been characterized as a "vertical mosaic" comprised of a unique mixture of several cultures which are loosely organized hierarchically along the lines of economics and political power (Porter, 1965, 1987). These cultures include the native peoples, or First Nations, who were the original inhabitants; Canadians of French descent who settled in Canada during the 17th and 18th centuries; and Canadians of British descent who came to Canada from Britain from the 17th century onwards or from the United States during the American Revolution in the latter part of the 18th century. The late 19th and 20th centuries brought three additional waves of immigration. In the latter part of the 19th century and early 20th century, immigrants from Asia, mainly China, came to the west coast of Canada. Immigrants from various European countries came to central Canada and the prairie provinces during the early 20th century and following World War II. More recently, Canada has experienced a new influx of immigrants and refugees, mostly from countries in economic or political turmoil: from Latin America, the Caribbean, Asia, Africa, the Indian subcontinent, the former Soviet Union, and the Middle East. All these cultural groups have had a major influence on Canadian society and its educational system.

PS 024022

BEST COPY AVAILABLE

Societal and Cultural Factors Affecting the Education of Children with Learning Disabilities

In this report we will address each of five societal and cultural factors which we have identified as important determinants of the education of children with learning disabilities in Canada. The first two relate to policy issues, the remaining three to issues of bilingualism and multiculturalism.

1. *Canadian attitudes toward public policy.* Canada has a relatively well-developed social safety net in that government funding supports extensive education, health, and social services. This has led to the expectation that services for children with learning disabilities should, for the most part, be provided by the public sector. Canadian attitudes toward education and service delivery have affected the process of program development for these children, and constitute an important factor differentiating Canada from the U.S.

2. *Provincial jurisdiction.* Jurisdiction of education in Canada is exclusively provincial (within a federal political system). Consequently, educational legislation and policy vary significantly from province to province.

3. *Official bilingualism.* Canada has two official languages—English and French—and this has important implications for students with learning disabilities who must either study or learn in a second language.

4. *Multiculturalism.* In addition to its English and French speaking populations, Canada has a large multicultural community. The dominant Canadian ethos holds that these communities should maintain the culture of their country of origin while simultaneously integrating into Canadian society. This approach to multiculturalism implies an acceptance and fostering of diversity that contrasts with U.S. policy on multiculturalism, which rests more explicitly on the individual's right to equal protection under the law as guaranteed by the U.S. Constitution. Diagnosing learning disabilities and differentiating them from problems stemming from the second language acquisition of English or French presents a major challenge.

5. *Native peoples.* Children from Native Canadian communities have specific language, learning, and cultural needs; 1.3% of Canada's children are from native groups.

Policy Issues

Canadian attitudes toward public policy. To characterize Canada's disposition toward the special education of children with learning disabilities, we rely here largely on our own observations. As one prominent Canadian sociologist has pointed out, there is a "paucity of information upon which to draw to construct a plausible Canadian character" (Porter, 1987, p. 90). It is our contention that Canadian attitudes have affected the making of policy and the mode of service delivery for children with learning disabilities.

Canadians tend to value and trust publicly funded institutions. They do not express the same kind of ambivalence about public and private responsibility for children's welfare that characterizes social attitudes in the U.S. (Jacobs & Davies, 1991). Canadians readily acknowledge the need for public provision of education, health, and social services, and they also prefer collaboration over litigation in the decision-making process. Lawsuits over educational matters, although they do occur, are rare in Canada. Moreover, a philosophy upholding the importance of the whole child pervades the Canadian approach to service provision.

How have these attitudes affected the development of services for children with learning disabilities? First, most service provision takes place within public institutions—through the school system primarily, but also in hospital and university affiliated learning centers and children's mental health clinics. Although some children may experience waits for services as long as 6 months, and more effective advocates get served more quickly, personal financial resources are seldom an issue. School boards and other agencies frequently make referrals to each other and collaborate on programs for individual children. Rarely, however, does one public institution (e.g., a school board) pay another (e.g., a hospital clinic) for services outside its jurisdiction. And distrust sometimes does arise between the private and public sectors.

Concurring with the literary critic and author Margaret Atwood, Porter writes that "survival is a dominant theme in the country's political life" and that our "concern for survival as a political entity places a premium on the practices of the past, [and] favors compromise and things as they are" (1987, p. 90). It is our view that Porter's interpretation, while not inaccurate, can be expressed more positively. Does a preference for compromise over confrontation, collaboration over litigation, really mean that "we Canadians are a conservative people" (1987, p. 90)?

One specialist writes that "service provision for the learning disabled essentially has been a nonadversarial process which has facilitated holistic, collaborative interventions involving educators, parents, peers, and most importantly, the learning disabled person." (Kronick, 1987, p. 1). It is our view that this preference for compromise and collaboration has benefited children with learning disabilities. School special services staff have been able to devote their time and energy to more informal types of assessment with the goal of developing instructional strategies rather than formal assessments aimed more narrowly at documenting need. Committee meetings to decide on placement and programming are usually informal and friendly. Even in provinces where actual legislation remains weak, services and programs have often been provided nonetheless, in the absence of mandatory policy. In addition, a prevailing holistic perspective has meant that at least some efforts have gone to developing social skills interventions, parent education, and counseling services.

This nonadversarial, publicly funded, and collaborative provision of services must be understood further in the context of the Canadian health care system, which funds clinics serving many children with learning disabilities. The health care system, unlike that of the U.S., is itself supported entirely by public funds and provides universal access. Public appreciation is widespread; litigation and malpractice suits are far less common, for instance, than in the U.S. For an excellent discussion of the Canadian health care system see "A search for solutions," *Consumer Reports*, 1992.

Provincial jurisdiction. Canada is comprised of ten provinces, the Northwest Territories, and the Yukon Territory. In 1988 the number of children in these provinces and territories ranged from 4,922 in the Yukon to 1,796,244 in Ontario. The four maritime provinces and the territories had a total child population of less than 500,000, with the Northwest Territories and Prince Edward Island having approximately 13,000 and 24,000 children respectively, and Newfoundland, Nova Scotia, and New Brunswick hovering around 150,000 each. Two of the three prairie provinces (Manitoba and Saskatchewan) had just over 200,000 children each. The child populations in Alberta, British Columbia, and Quebec were 426,476, 514,464, and 943,652, respectively.

Unlike the states of the U.S., the provinces of Canada have more autonomy from the federal government in the area of education. Each province has its own education act and policies pertaining to special education in general and learning disabilities in particular. All but three provinces—Prince Edward Island (the smallest province), Alberta, and British Columbia—have mandatory special education legislation (Poirier, Goguen, & Leslie, 1988). The legislation in these three provinces is permissive, that is, special education services may be provided by but not required of school boards.

The specific provisions of mandatory special education also vary from province to province (Poirier, Goguen, & Leslie, 1988). While mandatory legislation in the seven provinces that adopted it requires school boards to admit children with special needs, only in Manitoba and Quebec is this right to education universal, i.e., no child may be excluded from school. In the remaining five provinces the right to an education is quasi-universal in that some hard-to-serve children may be excluded. For example, in Ontario if a school board (the Canadian term roughly equivalent to a school district in the U.S.) determines that it does not have the facilities to provide for a specific child, the board may identify that child as "hard to serve." According to the 1980 amendments to the Education Act in Ontario, the school board's only obligation is to "assist the parent or guardian to locate a placement … suited to the needs of the pupil and reimburse the parent or guardian for any expenses incurred."

In only five provinces (Manitoba, New Brunswick, Ontario, Quebec, and Saskatchewan) are special education teachers required to have specialized certification and training. In only three provinces (Ontario, Quebec, and Saskatchewan) is there any provision that education must be appropriate to the needs of the child. Even in these three, formal approval of individual educational plans may not be required. Only in Quebec and Saskatchewan must children be placed in the least restrictive environment. The mandatory legislation in most provinces only covers school-aged children (ages 5 or 6, to 18 or 21); thus, in most provinces systematic early identification programs do not begin before school entry in kindergarten or grade 1. The exception is Quebec where 2,247 exceptional preschoolers are being served by the school system. In most other provinces children may be referred by parents, physicians, or others to hospital or community and social services clinics for evaluation and preschool programming. Finally, in only four provinces (Nova Scotia, Quebec, Ontario, and Saskatchewan) does the legislation state that parents have a right to be involved in the decision-making process.

Children with learning disabilities. Provinces differ in their provision of services for children with learning disabilities. In a survey conducted in 1988 by the Canadian Council for Exceptional Children (CEC), five of the ten provinces and both territories reported that they either did not categorize children at all or that children with mild disabilities received noncategorical services. It should be noted, however, that the five provinces that provide specific services for students with learning disabilities encompass 80% of the child population of Canada. One of the factors affecting the provision of services to children with learning disabilities is geography. Many of the provinces and territories providing only noncategorical services have small populations with a relatively high proportion of the population living in isolated rural communities (communities separated by more than 100 kilometers). Some experts have suggested that the financial and social costs of transporting children with learning disabilities to communities large enough to provide specialized services may outweigh the benefits (Bachor & Crealock, 1986).

The proportions of children formally identified as having learning disabilities varied widely, according to the CEC survey, in the five provinces that provide categorical service, with Quebec identifying 10.2% of its children, Nova Scotia 7%, Ontario 3.1%, Saskatchewan 1.7%, and British Columbia 1.3%. Identification rates within provinces may also show considerable variation. For example, a survey of the mental health needs of children and youth with learning disabilities in Metropolitan Toronto (Canada's largest city with a population of approximately 3 million people), identified 4.2% of school children as having learning disabilities, but the rates of the seven separate school boards ranged from 0.9% to 27% (Cummings, Hebb-Grier, Brazil, & Vallance, 1990). The definitions of learning disabilities are quite varied and not governed by formulae as are some in the U.S.

Services for children with learning disabilities also vary widely within the same province or adjacent school boards, but this is not mainly related to funding. Unlike school districts in the U.S., Canadian school boards do *not* receive provincial or federal funds based on the number of children identified with learning disabilities (or other handicaps). School board funds come from a combination of local property taxes and direct provincial funding, with provinces having formulae that adjust the level of provincial contribution based on the property tax base of the local community. There is no direct federal funding of education. Thus, school boards within provinces are relatively homogeneous in the amount of funds available to them. School boards in communities with high levels of poverty and large numbers of children with special needs may still find it difficult, however, to provide adequate services with the funds available.

The Constitution Act of 1982, which enacted the Canadian Charter of Rights and Freedoms, has had important implications for the rights of children with learning disabilities and other exceptional children. Although this federal law does not apply directly to education, the provisions of the Charter override virtually all provincial legislation. The equality rights provision of the Charter states:

> "Every individual is equal before and under the law and has the right to the equal protection and equal benefit of the law without discrimination and, in particular, without discrimination based on ... mental or physical disability" (Canadian Charter, 1982).

Of the four basic rights accorded every individual, "the right to equal benefit of the law" (i.e., a distribution of resources commensurate with need) is especially important (Henteleff,1990). The law suggests that even in the absence of mandatory special education legislation (as in Prince Edward Island and British Columbia) or where legislation has not specified appropriate or required programs, the Charter may provide this right. Recent cases brought before the Supreme Court, for example, have stipulated that parents have a right to a hearing of a case for special education for their child, even if provincial education legislation does not include this provision. Since no Charter case involving children with learning disabilities has yet reached the Supreme Court, there is considerable debate about how the Charter will be interpreted (Poirier, Goguen, & Leslie, 1988; Henteleff, 1990).

Bilingualism and Multiculturalism

Official bilingualism. English and French are the two official languages of Canada, and the federal government and the province of New Brunswick are officially bilingual. In Quebec, the official language is French and the majority of the population speaks French; in the rest of

Canada, the official language is English and the majority of the population speaks English. In most of Canada, alternative services in their native language are provided to minority Francophones (French-speaking people) or Anglophones (English-speaking people) where sufficient demand exists. Further, at some point in their education (the timing and amount varying within each province), children receive instruction in the second official language; approximately 50% of Anglophone children outside Quebec are receiving French language instruction at any given point in time (Statistics Canada, 1990), and all Quebec Francophone children receive English instruction from grade 4 through high school. In 1988-89, 1.9 million students were enrolled in second-language French courses. Also, approximately 5% of Anglophone children (228,000 children in 1988-89) are enrolled, at their parents' request, in bilingual or French immersion programs wherein at least half the school day (typically all but one hour of the school day) is spent studying in French.

What implications does official bilingualism have for children with learning disabilities? Lambert (1975) has used the terms "additive" and "subtractive" to describe the consequences of bilingualism. Positive or negative consequences are seen to be related to the context in which children are learning two languages, not to bilingualism per se. In some contexts, the learning of a second language can have either a positive impact or no impact on the development and maintenance of the child's mother tongue, and no negative consequences on cognitive or affective development. Lambert has called this *additive* bilingualism. In other contexts, the learning of a second language can have a negative impact on first language acquisition and possibly other negative consequences on cognitive and affective development. This is *subtractive* bilingualism. Assessing the additive or subtractive effects of bilingualism on children with learning disabilities may be particularly problematic.

In this analysis we examine three different contexts of language acquisition. The first involves the provision of second language instruction as a subject within the larger curriculum. Typically this "core" French or English program relies on decontextualized and academic methods of instruction whereby children study the language itself but do not use the language to study other subjects. Typically, the second language is not in common use in either the home or community. Children with learning disabilities may find core French or English instruction problematic because learning a second language in this way may be especially difficult for them (Bruck, 1982; Wiss, 1989). While core French or English instruction may not be subtractive for children with learning disabilities, in the sense that it detracts from the acquisition of their first language, it is possible that frequent failure experiences with the new language may have negative affective consequences; empirical evidence is lacking on this point, however. Such consequences (to the extent that they exist) may follow, however, from how the second language is taught. Second language instruction typically requires that students memorize verb forms, noun genders, and spelling, rather than that they converse and learn practical vocabulary and phrases.

A second context for language acquisition is that in which minority Anglophone or Francophone children receive instruction in their first language where this is *not* the dominant language of the province. This is the typical context for Anglophones in Quebec and Francophones outside Quebec. In some situations, such as that of Francophones in New Brunswick and Anglophones in Quebec, their first language is usually the dominant language of their family and local community and, in the case of Quebec Anglophones, the dominant language of North American culture and the media. These communities attempt to provide special education resources in the first language for children with learning disabilities.

Identifying children with learning disabilities within this context of bilingualism can prove difficult, however. Standardized tests normed for Quebec Francophones may not be appropriate for minority Francophones outside Quebec. Furthermore, even though legislation in provinces such as Ontario mandate it, recruiting qualified teachers to provide special education services in French to minority Francophone children with learning disabilities can be problematic. And even when teachers are themselves bilingual and trained in special education, they lack the methods with proven effectiveness to help children in these situations.

Francophones outside Quebec and New Brunswick usually represent such small minorities in their communities that they have to know English for daily living. In some cases French may cease to be the dominant language of the home, and a combination of English and French may become the child's first language. In some places where their numbers warrant it, such children may receive French education intended to maintain their language and culture. But according to Carey, "they are bilingual with incomplete learning of either language, and this provides for increased inconsistencies in sound to syllable mapping" (1987, p. 106). Similar to children with reading disabilities, they are slower in word naming and phonetic decoding, and have longer latencies for abstract as opposed to concrete words. Their educational attainment is lower than Anglophones. In short, the experience for many of these Francophone children is subtractive bilingualism (Carey, 1987).

A third context for the development of bilingualism includes French immersion and bilingual programs for Anglophone children. In French immersion programs, Anglophone children receive almost all of their education in French with only approximately one hour per school day devoted to English language arts. In bilingual programs, instruction is in English for half the day, in French the other half. Anglophone children typically enter immersion with well-developed skills in their first language but little exposure to French other than the short clips they see on Canadian *Sesame Street* (Carey, 1987). While French immersion programs may begin at any grade level, they begin most often in kindergarten or grade 1. Considerable evidence, including some from longitudinal studies, suggests that French immersion promotes additive bilingualism for the majority of Anglophone students (Carey, 1987; Lambert & Tucker, 1972). The extent to which this holds for children with learning disabilities has been the subject of considerable debate (Wiss, 1989). Some investigators assert that when children with learning disabilities are experiencing difficulty in both their French and English programs, it may be beneficial to maintain them in French immersion if they can obtain appropriate remedial assistance within the immersion program (Bruck, 1978; 1982; Cummins, 1979). Others suggest that a subgroup of children with learning disabilities who experience difficulty in early French immersion only, not in English, may benefit from being exempted from the immersion program (Trites, 1976, 1981).

A number of practical questions have arisen from the enrollment of children with learning disabilities in early French immersion programs. First, of course, is that at the time they enroll they have not yet received systematic reading instruction, and their learning disabilities are yet to be detected. Both researchers and practitioners question whether a reliable method can be devised to identify children with problems prior to school entry, before they encounter early immersion (Trites, 1986; Wiss, 1987).

How to diagnose learning disabilities once children have entered French immersion constitutes a second set of questions. Should they be assessed in English, in French, or in both? If achievement tests based on norms of children who are native speakers of English or French are deemed inappropriate—and the consensus is that they are (Carey, 1987; Wiss, 1987)—what tests or assessment methods should be used? If achievement is low in either or both languages, does this mean the child is necessarily learning disabled, or might he or she be simply responding poorly to immersion programming? At this point there are no answers to these questions. Children who experience difficulties are usually withdrawn from the French immersion program.

A third set of questions asks what the most appropriate placement and program should be once children with learning disabilities begin their schooling in French immersion programs. While children with learning disabilities will continue to experience learning difficulties when switched to an English program, are the problems as great as they would have been had they remained in French immersion (Bruck, 1978, 1982; Cummins, 1979)? Do different subgroups respond better to remaining in French immersion, or switching to English programming? Should special education be provided in French, English, or both? If special education in French is unavailable to children with learning disabilities in French immersion programs, as is frequently the case, should they remain in French immersion?

Official bilingualism has been a major political and social concern in Canada since confederation in 1867. This issue takes on special significance for children with learning disabilities, because the requirement that they learn a second language or learn in a second language may lead to additional frustration and failure. Whereas experiences with bilingual schooling may be typically additive for children without learning disabilities, the same experiences may subtract from the first language functioning of children with learning disabilities. These possible negative cognitive and affective consequences remain a subject of debate. Educational systems tend to discourage research on these issues, sometimes appearing reluctant to allow investigators permission to address these questions. Consequently, research that would contribute to this debate has been limited.

Multiculturalism. In addition to Anglophones and Francophones, Canada has a substantial population whose first language is neither English nor French.[1] The first language in this case is referred to as a "heritage language," and in Quebec the speakers are referred to as Allophones. Allophones in Canada (including native peoples), in 1986 (the last census), numbered 2,860,565, or 11.3% of the total population. Considering the recent wave of immigration, these numbers have likely increased since then. Allophone children are concentrated in the three largest metropolitan centers in Canada—Toronto, Montreal, and Vancouver—with more than 50% of children in Toronto and Vancouver coming from homes where neither English nor French is the native language. Except for a fairly large

---

[1] The proportion of non-native English speakers in Canada is much larger than in the U.S. In the U.S. 7.5% of residents over age 5 are native speakers of Spanish, and an additional 6.5% are native speakers of a language other than Spanish or English (Barringer, 1993). In Canada, in contrast, 40% of the population are native speakers of French, with an additional 11.3% native speakers of a language other than French or English (Statistics Canada, 1990).

Chinese community, most immigrants to Canada prior to 1970 came from Europe. More recent waves of immigrants and refugees have come from countries in the Caribbean and Latin America, from Africa, the middle East, the Indian subcontinent, and in particular, from Somalia, Vietnam, Sri Lanka, and Hong Kong.

Allophone children may receive instruction during the regular school program in their heritage language. Typically, heritage language programs involve 3 to 7 hours per week and take place in the school if there are sufficient numbers of children (usually at least 25) who speak that language. There are currently 82 heritage languages taught in the Toronto school system. In addition to programs in the schools, the government and the language communities themselves fund programs that provide instruction in the language and culture after school and/or on Saturdays.

Educational problems related to multiculturalism resemble to a certain extent those of official bilingualism, that is, how to facilitate additive bilingualism or trilingualism. There is universal agreement that Allophone children should learn at least one of the dominant languages (i.e., English or French). Educators also generally agree that strong functioning in the first language predicts good acquisition of a second and third language (Cummins, 1987). Where specialists do not agree is over the question of whether Allophone children who are weak in their first language should receive help to strengthen that language prior to or simultaneous with being taught a second or third language.

Allophone children may be weak in their first language for a variety of reasons (Cummins, 1984). Refugees may be victims of trauma and haphazard or nonexistent schooling. Others may have parents who neglect the child's first language because they believe that even though their own second language skills are limited they are helping their children by speaking to them in English or French. As a result the children receive poor language modeling in both languages. Finally, some Allophone children do indeed have learning disabilities.

Canadian researchers and educators face a challenging problem in diagnosing learning disabilities among Allophone children. Although most children readily acquire the basic interpersonal communication skills in a second language, linguistic proficiency in academic content can take five years or more (Carey, 1987; Cummins, 1984). Consequently it is often difficult to discriminate between those Allophone children with learning disabilities who may require special education and those children who simply need more instruction in English or French as a second language. One can conceptualize learning disabilities as a continuum where children at one end are clearly not learning disabled and those at the other end would have learning difficulties no matter what the educational environment. It has been suggested that along this continuum lies a group of children who are vulnerable, but who, under optimal conditions, would not develop learning disabilities (Adelman, 1989). Some Allophone students who are being asked to learn in a new language within a foreign culture may fit into this category.

Some researchers and educators have criticized Canadian school boards for excessive identification of Allophone children as learning disabled (Cummins, 1987). Consequently, progressive school boards currently use dynamic assessment techniques that involve teaching the child in the course of the assessment as a way of learning more about the child's skills, and they tend to be cautious about prematurely identifying Allophone children as learning disabled. This approach has the disadvantage that some Allophone children with learning disabilities who would benefit from special education services will not be identified. Such errors may be inevitable, however, until more reliable assessment methods are developed. Currently, investigators at McGill University in Montreal and the Ontario Institute for Studies in Education in Toronto are attempting to develop such tools.

A number of studies have found that participation in heritage language programs has no detrimental effect on second language acquisition, whether English or French, and that children in these programs also achieve reasonable proficiency in the heritage language (Cummins et al., 1984; da Fontoura & Siegel, in press; Swain, Lapkin, Rowen, & Hart, 1990; Siegel & So, 1993). Most of the heritage language groups in these studies performed as well in tests of language and reading proficiency as their monolingual, English-speaking peers. In addition, bilingual children who demonstrate reading problems in English have been shown to have similar problems in their heritage language, whether it is Portuguese (da Fontoura & Siegel, in press), Japanese (Saito & Siegel, 1993), Arabic (Abu Rabia & Siegel, 1993), or Cantonese (Siegel & So, 1993).

Native peoples. As already indicated, Canada's aboriginal or native peoples comprise approximately 1.3% of the population. Of this segment about 65% live in remote or rural communities, including the Inuit peoples, many of whom live in isolated regions in the North (Csapo, 1989). Although language plays a role in any discussion of learning disabilities within native groups (Native Canadians represent 10 language groups and 58 dialects), poverty and cultural differences present more pressing issues.

Native Canadians have suffered enormous exploitation, first by the European settlers, and subsequently by the dominant, majority governments; and individual groups have suffered assaults on their livelihood and culture. Poverty rates are high among native groups. Neonatal mortality is approximately 60% higher than the national rate, with the average death rate for native children three times that of the national rate (Csapo, 1989). The major causes of death among Native Indians are accidents, violence, and poisonings, with suicides being three times the national rate. It is estimated that 50% to 60% of Indian illnesses are alcohol-related, with fetal alcohol syndrome a relatively common problem among children. Due to the prevalence of persistent upper respiratory infections, children also frequently suffer hearing loss.

The education of Native Indians is the responsibility of the federal Department of Indian and Northern Affairs. Until the late 1960s many Natives received their education in residential schools whose aim was to eliminate their language and culture. Since 1970, however, the Department has run some schools on reserves and provides, in other cases, financing for schools to be run by the native bands themselves. Support is also provided for Natives to attend public schools under provincial jurisdiction. Nevertheless, researchers and advocates often express concerns about discrimination and the tracking of Natives into low academic streams.

The average educational attainment of native peoples is less than that of the non-native majority, with only 60% between the ages of 14 and 18 attending school, compared to a national average of 75%. University attendance is less than half the national average of 25%. In the larger context of physical health, mental health, and learning problems, the task of differentiating between low academic achievement that is due to poverty and cultural factors and that due to learning disabilities has not been a priority. It is clear, however, that native children are at risk, and that prevention programs are much needed.

Conclusion

Our discussion paints overall a promising picture of learning disabilities practice and research in Canada. But the analysis also reveals some problems. First, of concern is the fact that three provinces—British Columbia, Alberta, and Prince Edward Island—have not yet passed special education legislation mandating services. Although precedents under the federal Charter of Human Rights and Freedoms (Canadian Charter, 1982) suggest that exceptional children in these provinces do have a right to an education and parents to due process, the Charter does not set, and the courts have not required, specific standards for that education. Even the legislation in most provinces where quasi-universal or universal access has been established is somewhat deficient in the setting of such standards.

The lack of special education services for native children with learning disabilities constitutes a second problem. This lack of services, it must be noted, is not confined just to learning disabilities; native children also have serious economic, physical, and mental health needs requiring attention.

Although school-aged children with learning disabilities are relatively well provided for in Canada, serious gaps in services for preschool children persist. While there are many excellent diagnostic and treatment centers serving preschool children with learning and other developmental disabilities, few provinces have systematic screening programs for identifying them prior to school entry. Much more research needs to be conducted on appropriate assessments and remediation strategies for children with possible learning disabilities in a bilingual and multicultural setting. We do not yet know how best to identify these children or exactly how to treat their problems once identified.

Finally, Canada can nevertheless serve in some aspects as a model. As traditions have it, services for children with learning disabilities are relatively unfragmented and tend to serve the whole child. Teaching and teachers tend to receive respect. Collaborative consultation and collaborative advocacy are the rule, not the exception. Unlike in many states in the U.S., there is not usually the insistence on a strict definition of a learning disability in order for a child to be served. Consequently, in spite of legislative inadequacies and problems related to identification, most children with learning disabilities are able to receive services.

References

Abu Rabia, S., & Siegel, L. S. (1993). *Reading, language, and memory skills of bilingual Arabic-English speaking children*. Unpublished manuscript.

Adelman, H. S. (1989). Beyond the learning mystique: An interactional perspective on learning disabilities. *Journal of Learning Disabilities, 22*, 301-304.

Bachor, D. G., & Crealock, C. (1986). *Instructional strategies for students with special needs.* Toronto: Prentice Hall.

Barringer, F. (1993, April 28). Immigration in 80's made English a foreign language for millions. *New York Times*, p. A1, 10.

Bruck, M. (1978). The suitability of early French immersion programs for the language disabled child. *Canadian Journal of Education, 3*, 51-72.

Bruck, M. (1982). Language impaired children's performance in an additive bilingual education program. *Applied Psycholinguistics, 3*, 45-60.

Canadian Charter of Rights and Freedoms: Constitution Act. (1982). Schedule B of the *Canada Act* (U.K.) C.11.

Canadian Council for Exceptional Children. (1988). Status of educational service delivery to exceptional students: A provincial survey. *Keeping in Touch*, June.

Carey, S. T. (1987). Reading comprehension in first and second languages of immersion and Francophone students. *Canadian Journal for Exceptional Children, 3,* 103-108.

Csapo, M. (1989). *Children in distress: A Canadian perspective.* Vancouver, BC: Centre for Human Development Research.

Cummings, R. L., Hebb-Grier, A., Brazil, K., & Vallance, D. (1990). *Integra study of the mental health needs of children and youth with learning disabilities in metropolitan Toronto.* Toronto, ON: Integra Foundation.

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222-251.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy.* San Diego: College Hill Press.

Cummins, J. (1987). Psychoeducational assessment in multicultural school systems. *Canadian Journal for Exceptional Children, 3,* 115-117.

Cummins, J., Swain, M., Nakajime, K., Handscombe, J., Greens, D., & Tran, C. (1984). Linguistic interdependence among Japanese and Vietnamese immigrant students. In D. Rivera (Ed.), *Communicative competence: Approaches to language proficiency assessment, research and application* (pp. 60-81). London: Multilingual Matters Ltd.

da Fontoura, H. A., & Siegel, L. S. (in press). A study of the language and reading skills of bilingual Portuguese Canadian children. *Reading and Writing: An Interdisciplinary Journal.*

Henteleff, Y. M. (1990). *Recent judgements of the supreme court of Canada: New hope for the learning disabled.* Unpublished report submitted to the Learning Disabilities Association of Canada.

Jacobs, F. H., & Davies, M. W. (1991). Rhetoric or reality: Child and family policy in the United States. *Social Policy Report*, 5, 1-25.

Kronick, D. (1987). *A Linguistic/ conceptual/ dynamic model of social skills intervention.* Unpublished manuscript.

Lambert, W. E. (1975). Culture and language as factors in learning and education. In A. Wolfgang (Ed.), *Education of immigrant students.* Toronto: Ontario Institute for Studies in Education.

Lambert , W. E., & Tucker, R. (1972). *Bilingual education of children: The St. Lambert experiment.* Rowley: Newbury House.

Poirier, D., Goguen, L., & Leslie, P. (1988). *Educational rights of exceptional children in Canada: A national study of multi-level commitments.* Toronto: Carswell.

Porter, J. (1965). *The vertical mosaic: An analysis of social class and power in Canada.* Toronto: University of Toronto Press.

Porter, J. (1987). Canadian character in the twentieth century. In J. Porter (Ed.), *The measure of Canadian society: Education, equality and opportunity* (pp. 89-101). Ottawa: Carleton University Press.

A search for solutions (1992, September). *Consumer Reports, 57,* 579-592.

Saito, S., & Siegel, L. S. (1993). *The reading, language, and memory skills of Japanese Canadian children.* Unpublished manuscript.

Siegel, L. S., & So, D. (1993). *Reading, language, and memory skills of bilingual English Cantonese children.* Unpublished manuscript.

Statistics Canada. (1990). Minority and second language education: Elementary and secondary levels. Ottawa, ON: Author, 81-257.

Swain, M., Lapkin, S., Rowen, N., & Hart, D. (1990). The role of mother tongue literacy in third language learning. *Vox, 4*, 111-121.

Trites, R. (1976). Children with learning difficulties in primary French immersion. *Canadian Modern Language Review, 33*, 193-216.

Trites, R. (1981). Primary French immersion: Disabilities and prediction of success. *Review and Evaluation Bulletins 2*. Toronto: Ministry of Education, Ontario.

Trites, R. (1986). *Learning disabilities and prediction of success in primary French immersion: An overview*. Toronto: Ministry of Education, Ontario.

Wiss, C. (1987). Issues in the assessment of learning problems in children from French immersion programs: A case study illustration in support of Cummins. *The Canadian Modern Language Review, 43*, 302-313.

Wiss, C. (1989). Early French immersion programs may not be suitable for every child. *Canadian Modern Language Review, 45*, 517-529.

About the Authors

Linda Siegel is a professor and Judith Wiener is an associate professor at the Ontario Institute for Studies in Education, Toronto, Ontario.

Acknowledgements

The authors would like to thank Annette Hebb-Grier, Esther Geva, and Doreen Kronick for their valuable assistance in preparing and reviewing the manuscript. Preparation of this article was supported by a grant to L. S. Siegel from the Natural Sciences and Engineering Research Council of Canada and a grant to J. Wiener from the Social Sciences and Humanities Research Council of Canada. The article was prepared while L. S. Siegel held a Senior Research Fellowship from the Ontario Mental Health Foundation of Canada. The authors also wish to thank editor Nancy Thomas and reviewers Lynne Cook and Annette Zehler for their constructive suggestions.

Portions of this material were published in an article in the *Journal of Learning Disabilities*, 1992, *25*, 340-350 and are reprinted with permission.

Recent Issues

Volume V (1991)

No. 1. (Spring) *Two-generation program models: A new intervention strategy.*
Sheila Smith.

No. 2. (Summer) *Infant mortality and public policy.*
Ann L. Wilson and Gary Neidich.

No. 3. (Autumn) *The Migrant Head Start program.*
Mary Lou de Leon Siantz.

No. 4. (Winter) *Rhetoric or reality: Child and family policy in the United States.*
Francine H. Jacobs and Margery W. Davies.

Volume VI (1992)

No. 1. (Spring) *Mother-headed families: An international perspective and the case of Australia.*
Ailsa Burns.

No. 2. (Summer) *Testing in American schools: Issues for research and policy.*
Patricia Morison.

No. 3. (Fall) *The states and the poor: Child poverty rises as the safety net shrinks.*
Julie Strawn

No. 4 (Winter) *Crack's children: The consequences of maternal cocaine abuse.*
Theresa Lawton Hawley and Elizabeth R. Disney.

Volume VII (1993)

No. 2 (Summer) [Tentative] *Do research and theory justify Head Start's expansion?*
Edward Zigler and Sally J. Styfco.

Subscriptions ($10 to nonmembers of SRCD) and single issues ($3.50) are available from:

SRCD Executive Office
University of Michigan
300 North Ingalls, 10th floor
Ann Arbor, MI 48109-0406

Social Policy Report *is a quarterly publication of the Committee on Child Development, Public Policy and Public Information of the Society for Research in Child Development. Copyright of the articles published in the* Report *is maintained by the originator[s] of the work. Statements appearing in the* Report *are the views of the author and do not imply endorsement by the Editor or by SRCD.*

Editor: Nancy G. Thomas
Office: (303) 925-5516
Fax: (303) 925-9570

Announcements

- **Call for Papers**. Papers are invited for a special issue of *Children and Youth Services Review,* an International Multidisciplinary Review of the Welfare of Young People. Special issue editors, Professors Sandra K. Danziger and Sheldon Danziger, seek submissions of original research or program and policy analyses. Sample topics include, but are not limited to: antecedents and consequences of child poverty; determinants of trends in child poverty; effects of existing and proposed labor market, welfare, health, education, and community service programs and policies on the well-being of children and families; comparative or cross-national studies of child poverty and social policies; evaluation of demonstration projects and model programs for children and families. The deadline for submissions is September 15, 1993. For further information, contact Sandra K. Danziger, School of Social Work, 1065 Frieze Bldg., University of Michigan, Ann Arbor, MI 48109-1285.

- **Membership**. Division 37, the Division of Child, Youth, and Family Services of the American Psychological Association, welcomes membership inquiries. Division 37 focuses on an integration of knowledge of child and family development, clinical intervention, the design and implementation of service delivery systems, and the political process. The division includes a diverse group of clinicians, researchers, administrators, and advocates, and is not restricted to members of APA; student memberships are encouraged. For information, contact W. Rodney Hammond, Ph.D., School of Professional Psychology, Wright State University, Dayton, OH 45435.

SOCIAL POLICY REPORT
Society For Research In Child Development

Volume VII, Number 2 1993

USING RESEARCH AND THEORY TO JUSTIFY AND INFORM HEAD START EXPANSION

Edward Zigler
Sally J. Styfco

Head Start is America's favorite social program—or is it? President Clinton appeared to echo public sentiment in his first congressional address when he described the program as a sound social investment deserving of additional support. But what seemed like a unanimous call to fully fund Head Start was soon punctuated by voices of dissent. The attack should have been expected—the inevitable backlash that occurs whenever a favored position becomes so politically correct as to invite scrutiny. Also, conservatives were newly freed from their duty to side with the former Republican president, who vocally supported Head Start, and found the changing of the political guard an opportune time to counter whatever the Democratic victors held dear. They found ammunition in the constructive criticisms being offered by Head Start's friends, many of whom had silenced their concerns during prior administrations intent on reducing social welfare expenditures. Soon the print media began to fill with articles portraying Head Start as a national boondoggle, a "scam" that is a waste of precious tax dollars (Hood, 1992, 1993). The program's advocates felt their sense of victory turn into a defensive stance.

They have assumed this posture before. Since its inception, opinions of Head Start have bounced from positive to negative and back again. The project has been favored as a quick means of boosting poor children's IQ scores and criticized when these increases did not appear to be permanent. It has been praised as a cure for welfare dependency, teenage pregnancy, and criminality and condemned because no one ever collected data to substantiate such claims. Today Head Start is in the middle of an ideological debate over the proper role of government in solving social problems. Can the program help the nation attain its goal of having all children ready to learn when they begin school, or is it just another example of "Big Brother" intruding in the lives of private citizens?

Head Start has clearly been subjected to unrealistic expectations and entangled in the web of politics. The time is long overdue to look at this

PS 024022

program objectively, sifting fact from opinion and separating the emotions from the evidence that should be the primary basis for determining its future course. Head Start has been in operation for nearly 30 years and has been evaluated more extensively than perhaps any other social program. During this time the disciplines devoted to early childhood have blossomed and generated sophisticated theories and practical applications. The purpose of this paper is to use this accumulated wisdom to make a realistic assessment of Head Start. Such an analysis is vital not only to determining whether the program's record justifies its expansion. The knowledge derived from experience with Head Start can guide the program's evolution and inform the search for effective strategies of early intervention and education for children in poverty.

Head Start's Program and Goals

The current controversy over the value of Head Start is riddled with misconceptions about the program's structure and mission. Head Start is referred to either as child care, preschool education, health and social services, or a training and job program for poor adults. The project, in fact, serves all of these functions. Many observers also have confused Head Start with the Perry Preschool, an educational intervention project conducted in the 1960s and still being evaluated for lifetime effects. In reality, Head Start is a very different type of program than the Perry Project was, so the two should not so casually be compared. To enable constructive debate and effective social policy, interested parties must have an accurate view of Head Start's nature. This section describes Head Start's model, goals, and implementation to help clarify what this frequently misunderstood program actually encompasses.

A Brief History

Project Head Start originated during an optimistic period of American history, a time when there were both the desire and financial means to attempt to eliminate poverty and promote social equality. The federal government took a proactive role in this effort, launching the War on Poverty to provide education and self-help opportunities to enable socioeconomically disadvantaged citizens to improve their lives. The 1964 Economic Opportunity Act (EOA) opened the War by creating several projects for poor adults, including the Community Action Programs. The CAP were designed to assist local communities in establishing and administering their own antipoverty programs. However, some local governments opposed the CAP's proposed placement of administrative control and resources in the hands of impoverished citizens and therefore refused to apply for program grants. In an effort to make the CAP more palatable to local officials, while using what would have been an embarrassing budget surplus, the Head Start project was born. Sargent Shriver, head of the Office of Economic Opportunity and chief strategist of the Johnson Administration's War on Poverty, was struck by the realization that children comprised half of the population living below poverty. He envisioned a preschool intervention program as a way to "overcome a lot of hostility in our society against the poor," because no one could blame children for their financial predicament or fear giving them too much power (quoted in Zigler & Anderson, 1979, p. 12).

With the exception of a few compensatory education projects, there was very little experience at the time to suggest how to meet the needs of economically disadvantaged preschoolers. Issues of program content, size, and duration were therefore open questions for Shriver as he began piecing together a national intervention project.

His first step was to appoint a planning committee of 14 experts from the fields of preschool education, health, child development, and mental health. This diversity ensured that Head Start would become far more than an educational program. (For the history of the development of Head Start, see Zigler & Muenchow, 1992, and Zigler & Valentine, 1979.)

Dr. Robert Cooke, a noted pediatrician, chaired the group, which had only a few months to develop their plans. The committee's recommendations, presented to Shriver in February 1965, were based on a "whole child" philosophy that called for comprehensive programming: Head Start's goals were to improve physical health, enhance mental processes (particularly conceptual and verbal skills), and foster social and emotional development, self-confidence, relationships with family and others, social responsibility, and a sense of dignity and self-worth for both the child and family ("Recommendations for a Head Start Program," 1965). Note that only one of these seven goals specifically related to intellectual performance.

The committee's recommendations were unique not only because they outlined a multifaceted intervention but because they included the family as well as the child. The planning document accorded parents a central role that was virtually unprecedented in the design of preschool or compensatory education programs. Prior to Head Start, economically disadvantaged families had been treated as passive recipients of services dispensed by professionals. But Head Start parents were to be involved in the planning, administration, and daily activities of their local centers. One reason for this decision derived from Head Start's origins in the CAP, which were mandated to allow poor citizens "maximum feasible participation" in running antipoverty efforts. The EOA contained no clear definition of this slogan, and the Head Start planning committee interpreted it more literally than many local officials might have liked (see Valentine & Stark, 1979). The planners were also influenced by the ideas of one member, Urie Bronfenbrenner, who was just beginning to develop his ecological approach to human development. Bronfenbrenner (1974, 1979) argued that there is a complex interrelationship among children, their families, and communities, so intervention must touch all of these areas to be effective. This insight was an astute one: today, parent participation is recognized as crucial to the success of early intervention as well as later education (e.g., Comer, 1980; Powell, 1982; Seitz, 1990).

Program Implementation

Expert advisors and planning committee members counseled Shriver that a small, pilot program should be run and evaluated prior to mounting a large-scale effort. Yet the Johnson Administration demanded that Shriver fire a major volley in the War on Poverty by beginning on a large scale, a plan made possible by the economic prosperity of the times. When Head Start opened in 1965, over 500,000 children were served in a summer program lasting 6 or 8 weeks. Unfortunately, there was not enough time to screen all the grant applications carefully, so variation in quality characterized Head Start from the very beginning.

Today the basic Head Start program is a center-based preschool serving primarily poor children aged 3 to 5, with the majority (63%) being 4 years old (Administration on Children, Youth and Families, 1993). Most children attend a half-day session for a school year, although some participate for 2 years. Full-day programs and home-based services are delivered in some locations. Federal guidelines require that at least 90% of the children enrolled be from families whose income falls below the poverty line; at least

10% of enrollment must consist of children with disabilities. Head Start programs by law receive 80% of their funding from the federal government and the rest from other, usually local, sources, which may be in the form of in-kind services.

In fiscal year 1992, there were 1,370 Head Start grantees serving over 621,000 children and their families (Administration on Children, Youth and Families, 1993). (This is less than the yearly enrollment in the late 1960s but represents a 35% increase since 1989.) Each program is required to include six components: early childhood education, health screening and referral, mental health services, nutrition education and hot meals, social services for the child and family, and parent involvement. Local programs are overseen by Policy Advisory Councils, composed of parents, staff, and community representatives who are responsible for operating and staffing decisions. Although all Head Start programs must adhere to a body of national regulatory standards, each is encouraged to adapt the program to local needs and resources. Thus it is somewhat misleading to speak of Head Start as a singular intervention because of the variation in localized programming.

Evaluating Head Start

Head Start was designed as a comprehensive, two-generation, community-based program. Its complexity and uniqueness made evaluation difficult from the start. Federal officials understandably wanted a national impact study to ascertain the effects of this costly new experiment. Yet the diversity of Head Start programs around the country would confound comparisons of preschool and control groups, because the preschool samples would not all have had the same Head Start experience. The participants themselves were a heterogeneous population differing in terms of racial, ethnic, linguistic, geographic, and other characteristics. Further, random assignment to preschool and comparison groups would be problematic because eligibility for Head Start was defined by law. This left evaluation subject to less rigorous quasi-experimental designs (Cook & Campbell, 1979). Another problem was that there were no generally accepted definitions or measures of some concepts like social competence and parental empowerment that were part of Head Start's agenda.

The burst of research in the program's first decade illuminated these and other problems in evaluating such a complex intervention. Although this paved the way for more informed investigations, few were subsequently conducted. Support for Head Start research was greatly reduced during the 1980s, falling from 2.5% of the project's budget in 1974 to .11% in 1989, including demonstration efforts (National Head Start Association, 1990). Much of what is known about the effects of Head Start is therefore based on old, sometimes problematic studies, and, as we will discuss later in this section, there is much that is still not known.

Focus on Intelligence

Despite the many goals of Head Start, initial research focused almost exclusively on how much the program could raise children's intelligence test scores. This emphasis sprang from a popular view within developmental psychology that even limited environmental interventions could stimulate cognitive growth (see Schorr, 1988; Spitz, 1986; Zigler, 1970). Supporters of this position held that if the "right" experiences were provided at the "right" time (commonly assumed to be before age 5), IQ scores could be raised dramatically. Another reason that Head Start came to be judged on the basis of IQ changes could be that psychologists conducted most of the evaluations, and standard IQ tests are tools of their trade. If physicians had designed the first studies, perhaps Head Start's health benefits would have been evaluated instead. The positive impact of the health component (see below) might have become better known as a result.

Also contributing to the fixation on intellectual effects is the fact that the results were striking. Early studies showed that children's IQ scores leapt by at least 10 points and their achievement levels rose after just a few weeks of Head Start (see Datta, 1979). Then researchers at Ohio University and the Westinghouse Learning Corporation delivered the unwelcome news that the achievement gains faded away shortly after children entered school (Cicirelli, 1969). Experts in statistics and evaluation conducted in-depth reviews of the Westinghouse report and criticized it for problems with sampling procedures, data analysis, and appropriateness of the outcome measures (e.g., Campbell & Erlebacher, 1970; Datta, 1976; Lazar, 1981; Smith & Bissell, 1970). Nonetheless, subsequent studies of Head Start and almost every other early intervention program reached the same conclusion: preschool graduates generally do not continue to do better on cognitive tests or school quizzes. These reports dashed hopes that brief preschool experiences could guarantee academic success for poor children. Disappointed federal officials began to entertain plans to phase out Project Head Start.

Parents—empowered by their involvement in Head Start—came out in force and ultimately saved the program. Their efforts were soon bolstered by the release of findings from the Consortium for Longitudinal Studies (1978, 1983), a group of investigators who had evaluated 11 early intervention programs (2 of them Head Start) during the 1960s and early 1970s. The researchers located original program participants and collected a uniform set of information about them. Data from the programs were combined and analyzed using sophisticated techniques which avoided the problems that had weakened the validity of the Westinghouse findings. All significant results were reanalyzed by dropping the program with the strongest effect to determine whether it still held among the others (in effect, guaranteeing conservative conclusions).

The Consortium results confirmed that children who attend quality preschool programs evidence an initial boost in IQ and achievement scores that lasts for some years but eventually fades. However, lasting effects were found in other areas: participants were less likely to be assigned to special education classes, and were somewhat less likely to be held back a grade in school. The rigor of the Consortium methodology, and the findings of benefits that persisted until many children had reached 12 or more years of age, helped to restore faith in the value of early intervention.

Other longitudinal studies and reviews of research specific to Head Start programs have yielded similarly encouraging results. The Head Start Synthesis Project, a meta-analysis of some of over 200 studies reviewed, also found the loss of initial cognitive gains but reported that Head Start children had better health, immunization rates, and nutrition, as well as enhanced socioemotional characteristics (McKey et al., 1985). Family life was also found to be strengthened through the involvement of parents in the program. A study of thousands of children who had attended Head Start in 33 programs throughout Philadelphia showed that they had better school adjustment than peers who had no preschool experience (Copple, Cline, & Smith, 1987). After at least sixth grade, the preschool graduates had fewer school absences, did not miss as many standardized tests, and seemed less likely to be retained in grade. Another study of over 1,900 children who had attended Head Start at three different times revealed that by the end of high school, the oldest group performed better academically than a control sample (Hebbeler, 1985).

Although much of this research has technical problems that make the findings tentative, the results are all in the same direction: Head Start does not permanently raise IQ scores but may have enduring effects on school adjustment and other aspects of social competence (McCall, 1993). Findings concerning immediate program effects are much more definite: when children leave Head Start they have better IQ test scores and school readiness skills. In other words, Head Start does help prepare children for school. In lamenting the loss of initial cognitive benefits, many critics have overlooked Head Start's success in this regard.

There are several reasons why children leave the program testing well and then may lose this advantage. One is that performance on an IQ test reflects not only formal cognitive processes but achievement and personality variables as well. The IQ gains apparent after Head Start are not necessarily due to expanded intellectual capacity but may instead be explained by improved motivation, familiarity with test content, and comfort in the testing situation (Seitz, Abelson, Levine, & Zigler, 1975; Zigler, Abelson, Trickett, & Seitz, 1982; Zigler & Butterfield, 1968). Thus, participation in Head Start can enable children to develop the skills and attitudes needed to apply their abilities more fully. When they enter school, however, the environment may not continue to encourage full use of their potential; for example, there may be a poor curriculum or teaching practices that lower self-confidence and ignore individual learning styles. Another explanation for fade-out, advanced by early childhood educators, is that the paper-and-pencil, standardized group achievement tests used in many studies are inappropriate for children in the primary grades (e.g., Meisels, 1992). This argument supports Barnett's (1992) position that the achievement gains made in preschool probably do not fade out at all. He suggests that their apparent loss is an artifact of measurement, statistical analysis, and sampling procedures. For example, poor achievers in both preschool graduate and comparison groups are likely to be retained in grade and thus dropped from the evaluation, thereby minimizing the differences among the remaining students.

Confusion with Other Interventions

Many of the current criticisms of Head Start center on the argument that the project's results do not match those of more expensive preschool programs. This claim would not have assumed much importance if the findings from one such program—the Perry Preschool—did not become widely attributed to Head Start. Yet the fact that they did encourages us to examine not only the weaknesses in Head Start but the strengths, with the result being a more complete picture of the potential of early intervention.

The Perry Program was conducted by researchers at the High/Scope Educational Research Foundation in Ypsilanti, Michigan (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984). Like many other experimental projects, the program operated for only a few years and involved a relatively small number of children. What piqued scientists' interests was that the participants were followed into adulthood, with results reported for age 19 and recently for age 27 (Schweinhart, Barnes, & Weikart, 1993). What piqued public interest was a cost-benefit analysis which placed a monetary value on program effects.

The High/Scope investigators estimated the savings to society resulting from better social competence among Perry graduates, including reduced grade retention and usage of the welfare and criminal justice systems. They also counted higher contributions to the tax base resulting from higher employment rates. They reported that for every dollar spent on the preschool program, taxpayers received a savings of $3 to $6 by the time participants had reached 19 years of age (Barnett, 1985). By age 27, these savings were estimated to be $7 (Barnett, 1993).

The High/Scope researchers cautioned against generalizing their findings to the national Head Start program, a warning issued earlier by the Consortium investigators. The media, however, continued to devote attention to the Perry results, which were a lot more interesting and understandable than the positive but more esoteric reports on Head Start that were appearing in the scientific literature. Eventually the Perry Project came to be described as "Head Start-like," and its findings became attributed to early intervention in general and sometimes to Head Start itself.

The High/Scope program focused on preschool education, although there was a home visitation component to enable parents to reinforce the curriculum at home. Head Start also provides preschool education, but attempts to involve parents in many more ways and to deliver comprehensive services to children as well as their families. The Perry Project is therefore not at all "Head Start-like." And whether or not the two programs produce the same results is unknown. There has never been a report on Head Start's effects on college attendance or welfare usage, for example, and no one has ever conducted a cost-benefit analysis of the project (Haskins, 1989). Just considering the value of parental employment and preventive health care, Head Start may be more cost-effective than High/Scope. This is an empirical issue that must be addressed before comparative judgments of the two programs can be fairly made.

Despite the lack of comprehensive, longitudinal data on Head Start, many reviewers of the early intervention literature have accepted the assumption that Head Start graduates generally do not do as well as those who attended other, usually university-based programs (Haskins, 1989; Kotelchuck & Richmond, 1987; Woodhead, 1988). In defense of Head Start, some have pointed out that the smaller projects were very carefully designed and implemented, with highly trained staff and substantial budgets. The annual per-child cost of the Perry Preschool, for example, was estimated to be $7,600 in 1992 dollars (Barnett, 1993)—more than twice as much as Head Start spends. There is no way the national program could provide each child with more services with less than half the money. Further, the experimental preschools typically operated in single sites under constant monitoring to ensure that the programs were delivered in the manner intended. Head Start operates in nearly 1,400 different places where it is tailored to local circumstances and quality ranges from excellent to poor. Besharov has been particularly critical of the common assumption that Head Start produces the same benefits as other programs precisely because they do not have the same level of quality. Asserting that "not all preschool programs are created equal," Besharov believes that "Head Start has serious problems . . . that often prevent it from making a lasting impact on disadvantaged children" (1992, pp. 521, 520).

Because the experimental preschools focused almost exclusively on preschool education, the debate over generalizability of the evidence actually centers on the quality of the educational experience. The smaller programs typically contained an excellent curriculum delivered by well-trained teachers—features not found in all Head Start programs. The original planners purposefully avoided prescribing a national curriculum, so the quality of the educational component in each Head Start center essentially depends on local talent. It seems logical to assume that if all Head Start programs matched the educational standards of, for example, the Perry Preschool (including teacher qualifications and compensation), the same results would be achieved.

We would carry this assumption a step further and argue that with an improved preschool education component, Head Start's benefits would surpass those of the Perry Project. In addition to education, Head Start contains the other elements that many theorists and much empirical data associate with effective intervention: attention to the needs of the whole child, attention to the needs of the family through support services, and the involvement of parents, who are the child's first and most influential teachers (see National Head Start Association, 1990; Price, Cowen, Lorion, & Ramos-McKay, 1988; Schorr, 1988; Zigler & Berman, 1983; Zigler & Styfco, 1993). Even without this scientific support, it is logical to assume that a program that delivers quality educational and other needed services to children as well as their families will produce more benefits than one that provides a good education alone.

A question that begs answering is why the Perry preschoolers benefited from a program that had such a narrow focus when the knowledge base indicates that broader services are required. The High/Scope investigators believe that their sound curriculum allowed children to enter school with better academic skills and attitudes, which elicited positive reactions from their teachers; these further enhanced their performance and led to better student-teacher interactions in subsequent grades (Berrueta-Clement et al., 1984). This "snowball" hypothesis is consistent with Woodhead's suggestion that children who have quality preschool experiences are "better able to cope with the demands of schooling at a critical time when their identities within the education system [are] being established," triggering "a more positive cycle of achievement and expectation" (1988, p. 448). An alternative explanation centers on parents, who also have a critical role in helping their children adjust to school. As a result of 1 to 2 years of weekly home visits, the Perry parents may have become more interested and involved in their children's schooling (an outcome established for other programs, e.g., Hauser-Cram, Pierson, Walker, & Tivnan, 1991; Pfannenstiel, 1989). These attitudes, present at the sensitive time of transition to elementary school, could be what triggered the positive cycle of achievement. (See Hale, Seitz, & Zigler, 1990, and Seitz, 1990, for discussions that the High/Scope effects may be due to parents.)

Another question raised by the High/Scope results concerns the expectations we should have of early intervention. Recently, preschool education has been widely heralded as a means to end crime and welfare dependency. This ridiculous claim was certainly fueled by the publicity given the Perry findings. Yet the High/Scope investigators have pointed out that not all of their former students had good outcomes. Over half of the Perry graduates were arrested at least once by the time they were 27 years old, and 44% of the program females had been teen mothers (Schweinhart et al., 1993). Longitudinal studies of other preschool projects have also shown that although children do better than they would have without the experience, they still do not approach the achievements of middle-class students (e.g., Fuerst & Fuerst, 1993; Hebbeler, 1985).

Such findings lead to the sobering conclusion that early childhood intervention alone cannot transform lives. Its positive effects can be overpowered by the longer and larger experience of growing up in poverty. Quality preschool education such as that provided in the Perry Project may prepare poor children for school and help them meet social expectancies during the school years and perhaps beyond. Quality, comprehensive services such as those delivered in many Head

Start centers might achieve these and, as we are about to discuss, other important gains. But no program can enable children to develop optimally when their larger rearing environment is not conducive to optimal development. Early intervention can help, but it must not be considered more than one piece of the solution to the problem of poverty.

Lessons Yet to Learn from Head Start

Although the Perry Project was better than Head Start in terms of quality of preschool education (and in evaluation), in other respects Head Start is the superior program. The knowledge base clearly supports the value of Head Start's comprehensive, family-centered approach. Ironically, although the range of services provided in the Perry program was relatively narrow, the outcome measures used were broad, revealing the variety of behaviors affected by intervention. Of Head Start's many components, evaluators have devoted the most attention to the preschool aspect and its long-term effects on IQ and achievement. Possible benefits to physical health, nutritional status, social interactions, parents' child-rearing abilities, family functioning, parental empowerment, and community responsiveness have been underevaluated and undervalued.

The area in which Head Start has perhaps had its strongest impact is in physical health and well-being. Program performance standards require that all enrolled children receive medical screenings, immunizations, dental exams, and corrective treatment if needed. A high percentage of them do (Brush, Gaidurgis, & Best, 1993; National Head Start Association, 1990; Zigler, Piotrkowski, & Collins, in press), making Head Start a major provider of health services to poor children. This role is now expanding, as the project has recently been authorized to extend health care to participants' siblings. Head Start is also becoming a major supplier of screening and diagnostic testing services (EPSDT) required for children covered under Medicaid. In addition, Head Start delivers many of the services to preschoolers with disabilities and their families, as mandated under the Individuals with Disabilities Education Act (PL 99-457). In fact, children with disabilities comprise over 13% of Head Start's enrollment (Administration on Children, Youth and Families, 1993), so the program has done well in meeting the needs of this underserved population. Another health benefit derives from the nutritious meals and snacks served at each center.

Although Head Start's health benefits may seem obvious, little research has been done to document these effects. One study reported that Head Starters were just as likely as middle-class children to have inoculations and physical exams, and more likely to have seen a dentist (Hale et al., 1990). Similar physical exam histories were also found for the two groups' siblings. Despite the absence of more hard evidence, the value of Head Start's health services is really not in question. Each year hundreds of thousands of poor children receive some medical attention and nutritious food they otherwise might not have had. As the Clinton Administration searches for solutions to inadequate health care and dismal inoculation and hunger rates among children, the Head Start model may offer some guidance.

Head Start's positive impact on families is another outcome that has not received the attention it merits. In 1992, 94% of enrolled families received needed support services through the program (Brush et al., 1993). Each year thousands of low-income parents benefit financially by receiving jobs and job training through Head Start. Over 35% of the staff are parents of enrolled children or graduates (Brush et al., 1993), and many of them have earned Child Development Associate (CDA) credentials and have entered careers in early childhood education (Collins, 1990). Parents have also gained on a personal level, reporting improved relationships with their children (National Head Start Association, 1990) and greater life satisfaction and psychological well-being resulting from the supportive social network of the preschool community (Parker, Piotrkowski, & Peay, 1987).

Another possible effect on families is that siblings benefit from their parents' involvement in Head Start and the support services the family receives, as has been found for other interventions (reviewed by Seitz & Apfel, in press). It certainly makes sense that if parents are encouraged to be actively engaged in their children's education, and if they have help resolving at least some personal or practical problems, they may become better socializers of all their children. If so, the standard empirical approach that evaluates only the target child may greatly underestimate the benefits of Head Start.

Another underresearched area is the impact of Head Start on local communities. A few studies (reviewed by McKey et al., 1985) have shown that the presence of a Head Start program enhances a community's capacity to meet local needs. In one survey, investigators identified almost 1,500 institutional changes in the health care and educational systems in 48 communities after Head Start centers opened (Kirschner Associates, 1970). A more recent General Accounting Office (1992) report praised Head Start's methods of linking families with local services, judging this approach far more successful than efforts to create new services or delivery mechanisms. The GAO extended uncharacteristic approval by suggesting that "Congress may wish to consider promoting service-oriented efforts like Head Start" (p. 6).

Clearly, Head Start has the potential to enhance many systems that influence the course of a child's development. Judgments of the program based solely on academic indicators are therefore simplistic. To get a true picture of Head Start's effects on developmental outcomes, a broader empirical approach is required.

Policy Directions: Issues of Quantity and Quality

Despite its limitations, the literature on early childhood intervention points in two clear directions: to improve the life outcomes of children in poverty, programs must be more sustained and of high quality. We have learned that there are no magical periods of development when children can be inoculated against the ravages of social and economic disadvantage. Every period of development is important and requires appropriate environmental nutrients. Thus, to be more effective, intervention should begin earlier and last longer than the year or 2 before the child enters school. Further, not just any program will do. Only those that provide good services can be expected to achieve consistently good results. These lessons have important implications for the future of Head Start.

Dovetailed Programming

When Head Start began, a widely held belief was that a quick dose of environmental enrichment would immunize children against the effects of poverty. Fortunately, Head Start's planners did not subscribe to the inoculation model. They knew that no program could do much for a child in the course of one summer, which was the duration of Head Start in its first incarnation. Not only was the program extended to last for a school year, but planning got underway to serve children both before and after the preschool stage. Just 2 years after Head Start began, the first Parent and Child Centers (PCCs) were opened to provide supportive services and parent education to families and children from birth to age 3. That same year. Follow Through was initiated to continue Head Start services through the early grades of elementary school. Plans were for this to be a national project of the same scope as Head Start, but the escalating costs of the Vietnam War depleted the expected funding before the demonstration phase was completed (see Doernberger & Zigler, 1993). The concept survives as a tiny experiment in planned curriculum variation operating in less than 40 sites across the country. Sadly, after nearly 30 years, both of these efforts to extend Head Start have remained undeveloped and their potential gone unrealized.

There are a few encouraging signs. The number of PCCs has been expanded from 36 in 1989 to 106 today. The centers still operate without consistent performance standards, however, so there is no way of knowing what services are generally delivered or how well they meet the needs of participating families.

The first author has been working with the Carnegie Task Force on Meeting the Needs of Young Children to advise the Clinton administration on planning a 0 to 3 Head Start, possibly housed in existing PCCs. The exact details of the program would be worked out by a planning committee of experts from various disciplines similar to that which developed the current Head Start. Services for at-risk families would begin prenatally and include nutrition, health care, parenting education, and family support. (A child care component is also being suggested because of the dearth of quality infant care despite the growing need.) The rationale for the project is clear: waiting until a child is 3 or 4 years old is waiting too long. Preventive services can be more effective than remedial ones. Children who are healthy, have sound relationships with their primary caregivers, and who have received adequate nurturing and stimulation will have the socioemotional foundations needed for learning in preschool and beyond (Zero to Three, 1992).

A common complaint about Head Start is that graduates soon lose the academic and cognitive gains they made in preschool. This criticism is more appropriately directed toward the elementary schools they attend, for this is where the preschool advantage fades. To continue the momentum toward success, there must be a smooth transition from the preschool to the school environment and coordinated programming that builds upon prior learning.

The Head Start–Public School Early Childhood Transition Project (described by Kennedy, 1993) is the latest attempt to continue services in grade school. The project, which was legislated in the Human Services Reauthorization Act of 1990 and is now operating in 32 demonstration sites, follows Head Start graduates from kindergarten through grade 3. Local Head Start and public school personnel work to introduce each child and family to the new school experience and to familiarize kindergarten teachers with the child's progress, program, and needs. Comprehensive services, parental involvement, and family support are continued for the next 4 years. Early childhood educators often express skepticism about linking preschools and schools because they fear there will be a downward extension of formal educational practices that are not appropriate for very young children. In the Transition Project, curricula and pedagogies must be coordinated between the two levels of schooling. This should mean that the individualized, developmental nature of the preschool program will be extended upward, making the two school experiences less fragmented for young learners.

A small but convincing body of evidence supports the premise of the Transition Project, i.e., that longer, coordinated intervention produces longer lasting gains. When a very intensive intervention, the Abecedarian Project (Horacek, Ramey, Campbell, Hoffmann, & Fletcher, 1987), was followed by a school-age program through second grade, children continued to do better than peers who had not participated; classmates who experienced only the early intervention no longer had a statistically significant advantage over controls (Ramey & Campbell, 1991; Seitz, 1990), and those who received only the school-age program did not benefit substantially. An intensive reading curriculum in Success for All resulted in substantial gains for children who began the program in preschool, kindergarten, or first grade, and these gains magnified each year through third grade (Madden, Slavin, Karweit, Dolan, & Wasik, 1991). The results were much less pronounced for children who encountered the new curriculum after first grade, indicating that continuity from one reading level to the next was important for building program effects.

A program more similar to Head Start, the Chicago Child-Parent Centers, provided comprehensive services and required parental involvement for 1 to 2 years of preschool, kindergarten, and from 1 to 3 years of elementary school (Reynolds, 1992). Children who attended the preschool and kindergarten did not differ from controls by the time they were in fifth grade. However, fifth graders who had received 4 to 5 years of the intervention (preschool through second grade) had higher achievement scores and less grade retention, with no fade-out evident. Reynolds concluded that the transition between kindergarten and the primary grades is a crucial time, and that at least 2 years of follow-up programming in elementary school are required to sustain benefits.

A few studies suggest that the gains produced by extended intervention can last into high school and beyond. Initially, Abelson, Zigler, and DeBlasi (1974) found that Head Start graduates in their fourth year of Follow Through did better on IQ, achievement, and social-motivational measures than did their preschool classmates who attended traditional school programs. A follow-up showed these graduates maintained their superiority in several areas through grade 9 (Seitz, Apfel, Rosenbaum, & Zigler, 1983). In another study of the Chicago Child-Parent Centers, children who attended for 4 to 6 or more years had better high school graduation rates. For the overall sample, 62% received diplomas compared to about 49% of controls (Fuerst & Fuerst, 1993). Among students who had 7 to 9 years of intervention, 85% of girls and 70% of boys graduated. Center students who eventually dropped out did so 18 months later than other dropouts in the school system. Finally, a follow-up of children who had participated in the Deutschs' early enrichment program from preschool through third grade showed that by young adulthood, the boys had higher education and employment status and some stronger academic skills than a control sample (Jordan, Grallo, Deutsch, & Deutsch, 1985). The researchers attributed the loss of the advantage for girls to a more difficult transition to the regular fourth-grade classes, where their active learning orientations were not harmonious with traditional sex-role expectations.

This research has convinced us that the Transition Project will be effective in helping Head Start graduates maintain their preschool advantage. If evaluation of the demonstration phase confirms this expectation, it will be compelling to move the project into the educational mainstream. We have developed a plan to do so using current federal education expenditures (Zigler & Styfco, 1993). A large part of the Department of Education's budget (over $6 billion annually) is spent on Chapter 1 of the Elementary and Secondary Education Act of 1965 (ESEA). Chapter 1 is a compensatory education program for economically and "educationally" deprived children in preschool through grade 12. Originally intended to enhance the educational services of impoverished school districts, the program now operates in the majority of the nation's schools—mostly as a pull-out program offering remedial instruction to children who have fallen behind the academic expectations of their grade level.

There has been little evaluation of Chapter 1 considering the size of the program, but what there is shows that most students do not exhibit a meaningful gain in achievement (see Arroyo & Zigler, 1993). The reasons for the lackluster results may be in the program's design: services are remedial rather than preventive and narrow instead of comprehensive; parental involvement is minimal; and health and family problems that can interfere with school performance are not addressed. The ESEA will be reauthorized in 1993, and suggestions to revamp the program have focused on training teachers and narrowing the target population (Commission on Chapter 1, 1992). Although such efforts may do some good, they do not bring the elements of effective intervention to Chapter 1. To make a difference in the education of low-income children, we must put aside the ineffectual educational model of Chapter 1 and adopt on a large scale the proven model of comprehensive, family-focused services.

Our proposal is for Chapter 1 to follow Transition program plans and become the school-age version of Head Start. As Head Start eventually expands to serve all eligible children, Chapter 1 can continue their intervention in grammar school. Coordinated curricula and continued parental involvement and comprehensive services will then be firmly placed in schools that serve populations below the poverty line. On another level, Head Start and the Chapter 1 Transition will be two parts of a coherent federal policy to meet the needs of poor children beginning in preschool. Based on a solid knowledge base and big enough to have an impact, this new face for Chapter 1 holds promise for halting the fade-out of the Head Start advantage.

Renewed Emphasis on Quality

The studies involved in the Consortium, including the Perry Preschool, and others that were carefully designed, implemented, and evaluated (e.g., Price et al., 1988), prove the obvious: only good programs produce good results. For example, a review of early intervention projects led Weikart and Schweinhart to conclude that "only high quality programs consistently show success" (1991, p. 58). This fact may explain why some studies of Head Start show lasting benefits while others do not. One criticism of the Head Start Synthesis Project (McKey et al., 1985), which confirmed the fade-out observation, is that the analysts paid no attention to the differences in the Head Start programs in the studies they synthesized (Gamble & Zigler, 1989). Findings from studies of excellent programs were combined with those of poorer programs, diluting what may possibly have been robust effects for the former.

Quality problems are not new but have plagued Head Start since its hasty beginnings. In a matter of a few months, the project was transformed from an idea before the planning committee to a national summer program serving over one-half million children. Starting off so big and so fast left quality controls behind, and the program has been trying to catch up ever since. Program performance standards were not implemented until 1975. They dictate what services each program must provide and remain the principal vehicle for monitoring quality. The regional offices have primary responsibility for assuring that grantees adhere to the standards, but regional staffing has declined considerably over the years even as Head Start has become larger. One former regional director told us that his oversight staff decreased from at least 50 persons in 1978 to 12 in 1992, and travel money for site visits declined more than 90%. His experiences are not atypical; nationwide only one in five centers was monitored in 1988 (Chafel, 1992), although this level improved to 35% in 1992 (Horn, 1993). The best standards in the world are meaningless if they are not enforced.

Over the years many experts have studied the problems in Head Start and proposed thoughtful solutions. For the project's 15th anniversary the first author chaired a major advisory committee convened to chart future directions (U.S. Department of Health and Human Services, 1980). The report stressed areas where improvement was needed and made recommendations for corrective action. The committee's successor on the 25th anniversary, the Silver Ribbon Panel, found that many of the same problems were still there and had worsened over time (National Head Start Association, 1990). Other thorough analyses have similarly concluded that upgrading quality must be a top priority (e.g., Chafel, 1992; Washington & Oyemade, 1987; Zigler, Styfco, & Gilman, 1993).

Of all the program components, parental involvement and family support have been the most neglected. In many programs there are no coordinators for either component (or for health), and these positions are often combined with others because personnel budgets have fallen short of inflation (National Head Start Association, 1990). In 1991-92, social service staff had average caseloads higher than 94 children—more than three times the number recommended (Inspector General, 1993a). Between 1987 and 1991, training institutes were held in every area but parental involvement (National Head Start Association, 1991)—a void not addressed until 1993. Lack of attention to this component is particularly worrisome not only because participation has such potential benefit to parents (Nelson, 1991), but because they can support the intervention goals long after the formal program.

The biggest threat to Head Start quality is posed by staffing problems. Caregiver stability and level of skill have been identified as essential to quality care and directly responsible for positive developmental outcomes in children. In the National Child Care Staffing Study, quality of care was linked to staff wages and turnover rates (Whitebook, Howes, & Phillips, 1989). Low wages led to high turnover, disrupting continuity in the child's experiences and harming social and verbal development. In a nationwide survey of Head Start centers in 1988, Collins (1990) reported that teachers received an average salary of about $12,000, only 61% of the average compensation of teachers in public schools. In 1992, starting salaries for Head Start teachers averaged only $12,077 (Brush et al., 1993), about half the public school average. Despite poor salaries, Head Start has been spared from high attrition apparently because staff are committed to their jobs. Collins found annual turnover in Head Start to be about 25%, considerably lower than that in other early care and education programs. Yet a more recent report showed that in Head Start centers that do experience turnover, the rate is higher than 60% (Kisker, Hofferth, Phillips, & Farquhar, 1991). Quality care cannot be provided under such conditions.

Critics are quick to note that funding could not possibly be the cause of Head Start's problems because its budget has skyrocketed, jumping from $1.2 billion in 1989 to $2.8 billion in 1993. But if we look at the numbers prior to the recent expansion increases, we can see why quality has sometimes been compromised. Head Start barely held its funding level until the positive results of the Consortium studies were released. Then the project received its first substantial budget increase and was placed in a "safety net" where it would be spared spending reductions the Reagan Administration imposed on most federal programs. However, cuts in federal food and other programs that centers relied upon began to erode their ability to maintain services even at previous levels (Washington, 1985). Despite increasing appropriations, the average expenditure per child fell by more than $400 in constant dollars, or 13% in inflation-adjusted dollars, between 1981 and 1989 (Rovner, 1990). It is no wonder that a recent Inspector General report (1993a) showed that in some Head Start programs, participants do not receive even basic services such as all their immunizations. Despite questions about the criteria used in the report, and indeed whether the centers surveyed were representative, there is no doubt that some Head Start centers deliver inadequate services.

Years of insufficient funding are of course not the only reason Head Start has been struggling to maintain quality. The population it serves and their needs have changed. America is a different place than it was in 1965. The face of poverty has grown uglier. The environments where poor children are raised now involve more homelessness, street violence, illegal drugs, and young, single-parent families. Children who were born with the HIV virus or exposed to crack are now showing up in Head Start centers. The number of dysfunctional families with myriad needs has multiplied. The extra attention required by these children and families has burdened already overworked staff and undoubtedly deprived some participants of the individual help they need (Inspector General, 1989, 1993b).

Quality can no longer be threatened or overlooked, especially now that Head Start is serving increasing numbers of children and families. The Silver Ribbon Panel warned that "Program expansion should *never* occur at the expense of quality" (National Head Start Association, 1990, p. 35). Supporting their conclusion, the first author testified before the U.S. Senate (1990, p. 49) that "Head Start is effective only when quality is high. . . . Below a certain threshold of quality, the program is useless, a waste of money regardless of how many children are enrolled." Policymakers have good reasons to want to open Head Start to every eligible child, but they must assure each of them a good Head Start experience.

Planning for a Better, Bigger Program

Lawmakers heeded the advice of experts when they wrote a quality set-aside into the Human Services Reauthorization Act of 1990. After adjustments for inflation, 25% of expansion funds are to be used for improvements. Half of the set-aside is reserved for increased salaries and benefits. The rest is marked for training and technical assistance, facility improvements, and transportation. The Act also requires that by 1994, each center is to have at least one teacher with a CDA credential. While a step in the right direction, this mandate shows how far Head Start has fallen behind over the years and how far it has to go to catch up with current professional practices. Few middle-class parents would enroll their children in a preschool program where the teachers did not have college degrees and certification in early childhood education. Head Start is supposed to be a model program and should also have qualified teachers who are paid a worthy wage.

Although the plans to improve Head Start are good ones, they cannot be implemented overnight. Training staff and locating facilities, for example, can be very time consuming. But program directors are also under a great deal of pressure to serve more children, which can drain the energies needed for trying to serve them better.

Another Inspector General's report (1993b) and discussions with program directors tell the story. In some areas Head Start centers have been unable to recruit all the children they are supposed to enroll with expansion monies. A widespread problem is the inability to find qualified staff because they are either unavailable or will not hire on at the low pay levels. Space is a major problem, with some centers having to operate with double sessions because new classrooms cannot be found. This practice reduces time for planning daily programs and attending to family needs. Many programs have had a difficult time raising the local match to their sudden budget increases.

Expansion is clearly proceeding without a plan. The effort is of such magnitude and importance that one must be implemented quickly so that growth does not threaten efforts to enhance quality. Thankfully, both the White House and Donna Shalala, Secretary of Health and Human Services (HHS), have voiced a commitment to thoughtful growth *and* improving quality. Shalala appointed the Advisory Committee on Head Start Quality and Expansion, charged with recommending the priorities and pace for expansion. Their work also focused on short- and long-term plans to strengthen services. The Committee's recommendations, now in various stages of review and implementation within Congress and the agencies responsible for Head Start, have at the very least begun the planning process in earnest.

Although much attention is now being devoted to poor service delivery in some Head Start centers, expansion and improvement efforts must apply to administration as well. Some administrative responsibilities have been scattered throughout HHS and should be brought back to the central Head Start office. For example, staffing for health care was moved to the Maternal and Child Health Division, leaving a void in leadership for this vital program component. Regional Head Start personnel no longer report directly to Head Start's top managers, so there is not a clear managerial relationship between the national headquarters and those who are responsible for carrying out the program. And, as noted earlier, regional staffing levels have been greatly reduced over the years, so there are fewer people to oversee more grantees. To have a better Head Start program, we must have a well-managed one that is sufficiently staffed to accomplish its charge.

Plans to improve Head Start must also address the future. The project's success rests on its capacity to meet the needs of impoverished children and their families; as their needs change, so must the program. The Advisory Panel for the Head Start Evaluation Design Project (1990) presented a blueprint for needed research to guide the development of better services and delivery mechanisms. Their recommendations focused on studying which program services work best and for whom. This overriding question suggests several circumscribed ones: What mediates program effects? Are 2 years of Head Start more or less cost-effective than 1? Do all participants need all program services?

Another question—one that can no longer be answered by generalizing from the findings of other programs—is just how wide-ranging and lasting are Head Start's benefits? In 1990 Congress entertained plans to conduct a 20-year longitudinal study to include competence indicators such as rates of criminality, welfare use, and employment status of Head Start graduates and their families. Unfortunately, this much needed work was never funded. Hopefully the current controversy over the value of Head Start, raging at a time when substantial monies are being granted for expansion, will resurrect the commitment to research.

This is not an empty hope. After 30 years of experience with early intervention, good ideas and plans for empirical study have been developed. A committee charged with implementing the Design Project's blueprint is now in session. And Secretary Shalala has repeatedly acknowledged the need for scientific guidance to enhance program effectiveness. Head Start is finally moving toward long-awaited expansion and improvement of service quality. These efforts must be anchored in the best knowledge available if Head Start is to fulfill its promise of improving the lives of children and families in poverty.

References

Abelson, W. D., Zigler, E., & DeBlasi, C. L. (1974). Effects of a four-year Follow Through program on economically disadvantaged children. *Journal of Educational Psychology, 66*, 756-771.

Administration on Children, Youth and Families. (1993, January). *Project Head Start. Statistical Fact Sheet.* Washington, DC: U.S. Department of Health and Human Services.

Advisory Panel for the Head Start Evaluation Design Project. (1990). *Head Start research and evaluation: A blueprint for the future.* Washington, DC: Administration on Children, Youth and Families.

Arroyo, C. G., & Zigler, E. (1993). America's Title I/Chapter 1 programs: Why the promise has not been met. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond: A national plan for extended childhood intervention* (pp. 73-95). New Haven, CT: Yale University Press.

Barnett, W. S. (1985). The Perry Preschool Program and its long-term effects: A benefit-cost analysis. *High/Scope Early Childhood Policy Papers* (No. 2). Ypsilanti, MI: High/Scope Educational Research Foundation.

Barnett, W. S. (1992). Benefits of compensatory preschool education. *Journal of Human Resources, 27*, 279-312.

Barnett, W. S. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal of Orthopsychiatry, 63*, 500-508.

Berrueta-Clement, J. R., Schweinhart, L., Barnett, W., Epstein, A., & Weikart, D. (1984). *Changed lives: The effects of the Perry Preschool Program on youths through age 19.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Besharov, D. J. (1992). New directions for Head Start. *The world and I, 7*, 515-531.

Bronfenbrenner, U. (1974). Is early intervention effective? *Day Care and Early Education, 44*, 14-18.

Bronfenbrenner, U. (1979). *The ecology of human development.* Cambridge, MA: Harvard University Press.

Brush, L., Gaidurgis, A., & Best, C. (1993). *Indices of Head Start Program Quality.* Report prepared for the Administration on Children, Youth and Families, Head Start Bureau. Washington, DC: Pelavin Associates.

Campbell, D. T., & Erlebacher, A. (1970). How regression artifacts in quasi-experimental evaluations can mistakenly make compensatory education look harmful. In J. Hellmuth (Ed.), *Compensatory education: A national debate* (Vol. 3, pp. 185-210). New York: Brunner/Mazel.

Chafel, J. A. (1992). Funding Head Start: What are the issues? *American Journal of Orthopsychiatry, 62*, 9-21.

Cicirelli, V. G. (1969). *The impact of Head Start: An evaluation of the effects of Head Start on children's cognitive and affective development.* Report presented to the Office of Economic Opportunity (Report No. PB 184 328), Westinghouse Learning Corporation. Washington, DC.

Collins, R. C. (1990). *Head Start salaries: 1989-90 staff salary survey.* Alexandria, VA: National Head Start Association.

Comer, J. P. (1980). *School power.* New York: Free Press.

Commission on Chapter 1. (1992). *Making schools work for children in poverty.* Washington, DC: U.S. Department of Education.

Consortium for Longitudinal Studies. (1978). *Lasting effects after preschool.* Final report to Administration on Children, Youth and Families (Pub. No. OHDS 79-30178). Washington, DC: U.S. Government Printing Office.

Consortium for Longitudinal Studies. (Ed.). (1983). *As the twig is bent: Lasting effects of preschool programs.* Hillsdale, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Copple, C., Cline, M., & Smith, A. (1987). *Paths to the future: Long-term effects of Head Start in the Philadelphia school district.* Washington, DC: U.S. Department of Health and Human Services.

Datta, L. (1976). The impact of the Westinghouse/ Ohio evaluation on the development of Project Head Start. In C. C. Abt (Ed.), *The evaluation of social programs* (pp. 129-181). Beverly Hills, CA: Sage.

Datta, L. (1979). Another spring and other hopes: Some findings from national evaluations of Project Head Start. In E. Zigler & J. Valentine (Eds.), *Project Head Start: A legacy of the War on Poverty* (pp. 405-432). New York: Free Press.

Doernberger, C., & Zigler, E. (1993). Project Follow Through: Intent and reality. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond: A national plan for extended childhood intervention* (pp. 43-72). New Haven, CT: Yale University Press.

Fuerst, J. S., & Fuerst, D. (1993). Chicago experience with an early childhood program: The special case of the Child Parent Center program. *Urban Education, 28*, 69-96.

Gamble, T., & Zigler, E. (1989). The Head Start Synthesis Project: A critique. *Journal of Applied Developmental Psychology, 10*, 267-274.

General Accounting Office. (1992). *Integrating human services* (Report No. GAO/HRD-92-108). Washington, DC: Author.

Hale, B. A., Seitz, V., & Zigler, E. (1990). Health services and Head Start: A forgotten formula. *Journal of Applied Developmental Psychology, 11*, 447-458.

Haskins, R. (1989). Beyond metaphor. The efficacy of early childhood education. *American Psychologist, 44*, 274-282.

Hauser-Cram, P., Pierson, D. E., Walker, D. K., & Tivnan, T. (1991). *Early education in the public schools.* San Francisco: Jossey-Bass.

Hebbeler, K. (1985). An old and a new question on the effects of early education for children from low income families. *Educational Evaluation and Policy Analysis, 7*, 207-216.

Hood, J. (1992, December). Caveat emptor: The Head Start scam. *Policy Analysis*, #187. Washington, DC: Cato Institute.

Hood, J. (1993, February 19). What's wrong with Head Start. *Wall Street Journal*, p. A14.

Horacek, H., Ramey, C., Campbell, F., Hoffmann, R., & Fletcher, R. (1987). Predicting school failure and assessing early intervention with high-risk children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 758-763.

Horn, W. F. (1993). *Administration on Children, Youth and Families. Accomplishments: FY 1989 to FY 1993.* Washington, DC: ACYF.

Inspector General (Office of). (1989). *Dysfunctional families in the Head Start program: Meeting the challenge.* Washington, DC: U.S. Government Printing Office.

Inspector General (Office of). (1993a). *Evaluating Head Start expansion through performance indicators* (OEI-09-91-00762). Washington, DC: U.S. Department of Health and Human Services.

Inspector General (Office of). (1993b). *Head Start expansion: Grantee experiences* (OEI-09-91-00760). Washington, DC: U.S. Department of Health and Human Services.

Jordan, T. J., Grallo, R., Deutsch, M., & Deutsch, C. P. (1985). Long-term effects of early enrichment: A 20-year perspective on persistence and change. *American Journal of Community Psychology, 13*, 393-415.

Kennedy, E. M. (1993). The Head Start Transition Project: Head Start goes to elementary school. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond: A national plan for extended childhood intervention* (pp. 97-109). New Haven, CT: Yale University Press.

Kirschner Associates. (1970). *A national survey of the impacts of Head Start centers on community institutions.* Albuquerque, NM: Author.

Kisker, E. E., Hofferth, S. L., Phillips, D. A., & Farquhar, E. (1991). *A profile of child care settings: Early care and education* (Vol. 1). Washington, DC: U.S. Department of Education.

Kotelchuck, M., & Richmond, J. B. (1987). Head Start: Evolution of a successful comprehensive child development program. *Pediatrics, 79*, 441-445.

Lazar, I. (1981). Early intervention is effective. *Educational Leadership, 38*, 303-305.

Madden, N. A., Slavin, R., Karweit, N., Dolan, L., & Wasik, B. (1991). Success for all. *Phi Delta Kappan, 72*, 593-599.

McCall, R. (1993). *Head Start: Its potential, its achievements, its future.* A briefing paper for policymakers. Pittsburgh, PA: University of Pittsburgh Center for Social and Urban Research.

McKey, R. H., Condelli, L., Ganson, H., Barrett, B., McConkey, C., & Plantz, M. (1985). *The impact of Head Start on children, family, and communities: Final report of the Head Start Evaluation, Synthesis and Utilization Project* (DHHS Pub. No. OHDS 85-31193). Washington, DC: U.S. Government Printing Office.

Meisels, S. J. (1992). Doing harm by good: Iatrogenic effects of early childhood enrollment and promotion policies. *Early Childhood Research Quarterly, 7*, 155-174.

National Head Start Association. (1990). *Head Start: The nation's pride, a nation's challenge.* Report of the Silver Ribbon Panel. Alexandria, VA: Author.

National Head Start Association. (1991). Parent involvement survey. *NHSA Journal, 10*(2), 30-31.

Nelson, T. (1991). *What really makes Head Start work? You may be surprised.* Milwaukee, WI: Institute for Innovation in Human Services.

Parker, F. L., Piotrkowski, C. S., & Peay, L. (1987). Head Start as a social support for mothers: The psychological benefits of involvement. *American Journal of Orthopsychiatry, 57*, 220-233.

Pfannenstiel, J. (1989). *New Parents As Teachers project: A follow-up investigation.* Jefferson City: Missouri Department of Elementary and Secondary Education.

Powell, D. R. (1982). From child to parent: Changing conceptions of early childhood intervention. *Annals of the American Academy of Political and Social Science, 461*, 135-144.

Price, R. H., Cowen, E., Lorion, R. P., & Ramos-McKay, J. (Eds.). (1988). *Fourteen ounces of prevention: A casebook for practitioners.* Washington, DC: American Psychological Association.

Ramey, C. T., & Campbell, F. A. (1991). Poverty, early childhood education, and academic competence: The Abecedarian experience. In A. C. Huston (Ed.), *Children in poverty. Child development and public policy* (pp. 190-221). New York: Cambridge University Press.

Recommendations for a Head Start Program by a Panel of Experts. (1965, February). Washington, DC: U.S. Department of Health, Education and Welfare, Office of Child Development.

Reynolds, A. J. (1992). *Effects of a multi-year child-parent center intervention program on children at risk.* Manuscript submitted for publication.

Rovner, J. (1990). Head Start is one program everyone wants to help. *Congressional Quarterly, 48*(16), 1191-1195.

Schorr, L. B. (1988). *Within our reach: Breaking the cycle of disadvantage.* New York: Doubleday.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Seitz, V. (1990). Intervention programs for impoverished children: A comparison of educational and family support models. *Annals of Child Development, 7*, 73-103.

Seitz, V., Abelson, W. D., Levine, E., & Zigler, E. (1975). Effects of place of testing on the Peabody Picture Vocabulary Test scores of disadvantaged Head Start and non-Head Start children. *Child Development, 46*, 481-486.

Seitz, V., & Apfel, N. H. (in press). Parent-focused intervention: Diffusion effects on siblings. *Child Development.*

Seitz, V., Apfel, N., Rosenbaum, L., & Zigler, E. (1983). Long-term effects of Projects Head Start and Follow Through: The New Haven Project. In Consortium for Longitudinal Studies (Eds.), *As the twig is bent: Lasting effects of preschool programs* (pp. 299-332). Hillsdale, NJ: Erlbaum.

Smith, M., & Bissell, J. S. (1970). Report analysis: The impact of Head Start. *Harvard Educational Review, 40*, 51-104.

Spitz, H. H. (1986). *The raising of intelligence: A selected history of attempts to raise retarded intelligence.* Hillsdale, NJ: Erlbaum.

U.S. Department of Health and Human Services. (1980). *Head Start in the 1980's. Review and recommendations.* Washington, DC: Author.

U.S. Senate. (1990, August 3). *Human Services Reauthorization Act of 1990. Report to accompany H.R. 4151* (Report 101-421). Washington, DC.

Valentine, J., & Stark, E. (1979). The social context of parent involvement in Head Start. In E. Zigler & J. Valentine (Eds.), *Project Head Start: A legacy of the War on Poverty* (pp. 291-313). New York: Free Press.

Washington, V. (1985). Head Start: How appropriate for minority families in the 1980s? *American Journal of Orthopsychiatry, 55,* 577-590.

Washington, V., & Oyemade, U. (1987). *Project Head Start: Past, present, and future trends in the context of family needs.* New York: Garland.

Weikart, D. P., & Schweinhart, L. J. (1991). Disadvantaged children and curriculum effects. *New Directions for Child Development, 53,* 57-64.

Whitebook, M., Howes, C., & Phillips, D. (1989). *Who cares? Child care teachers and the quality of care in America.* Final report, National Child Care Staffing Study. Oakland, CA: Child Care Employee Project.

Woodhead, M. (1988). When psychology informs public policy. The case of early childhood intervention. *American Psychologist, 43,* 443-454.

Zero to Three: National Center for Clinical Infant Programs. (1992). *Heart start: The emotional foundations of school readiness.* Arlington, VA: Author.

Zigler, E. (1970). The environmental mystique: Training the intellect versus development of the child. *Childhood Education, 46,* 402-412.

Zigler, E., Abelson, W. D., Trickett, P. K., & Seitz, V. (1982). Is an intervention program really necessary to raise disadvantaged children's IQ scores? *Child Development, 53,* 340-348.

Zigler, E., & Anderson, K. (1979). An idea whose time had come: The intellectual and political climate for Head Start. In E. Zigler & J. Valentine (Eds.), *Project Head Start: A legacy of the War on Poverty* (pp. 3-19). New York: Free Press.

Zigler, E., & Berman, W. (1983). Discerning the future of early childhood intervention. *American Psychologist, 38,* 894-906.

Zigler, E., & Butterfield, E. C. (1968). Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. *Child Development, 39,* 1-14.

Zigler, E., & Muenchow, S. (1992). *Head Start: The inside story of America's most successful educational experiment.* New York: Basic Books.

Zigler, E., Piotrkowski, C., & Collins, R. (in press). Health services in Head Start. *Annual Review of Public Health.*

Zigler, E., & Styfco, S. J. (1993). Strength in unity: Consolidating federal education programs for young children. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond: A national plan for extended childhood intervention* (pp. 111-145). New Haven, CT: Yale University Press.

Zigler, E., Styfco, S. J., & Gilman, E. (1993). The national Head Start program for disadvantaged preschoolers. In E. Zigler & S. J. Styfco (Eds.), *Head Start and beyond: A national plan for extended childhood intervention* (pp. 1-41). New Haven, CT: Yale University Press.

Zigler, E., & Valentine, J. (Eds.). (1979). *Project Head Start: A legacy of the War on Poverty.* New York: Free Press.

About the Authors

Edward Zigler is Sterling Professor of Psychology at Yale University, where he is Director of the Bush Center in Child Development and Social Policy. He was a member of the original planning committee for Head Start and was the federal official responsible for administering the program from 1970-72, when he was Director of the Office of Child Development (now the Administration on Children, Youth and Families). Most recently he served on the Advisory Committee on Head Start Quality and Expansion and is a member of the Advisory Panel on Implementing the Head Start Blueprint for the Future.

Sally J. Styfco is a writer at the Yale University Bush Center in Child Development and Social Policy, where she is a specialist in educational policy issues affecting children and their families.

Acknowledgments

The authors thank Saul Rosoff and Sheldon White for their helpful comments on an earlier draft of the manuscript.

Recent Issues

Volume V (1991)

No. 1. (Spring) *Two-generation program models: A new intervention strategy.*
Sheila Smith.

No. 2. (Summer) *Infant mortality and public policy.*
Ann L. Wilson and Gary Neidich.

No. 3. (Autumn) *The Migrant Head Start program.*
Mary Lou de Leon Siantz.

No. 4. (Winter) *Rhetoric or reality: Child and family policy in the United States.*
Francine H. Jacobs and Margery W. Davies.

Volume VI (1992)

No. 1. (Spring) *Mother-headed families: An international perspective and the case of Australia.*
Ailsa Burns.

No. 2. (Summer) *Testing in American Schools: Issues for research and policy.*
Patricia Morison.

No. 3. (Fall) *The states and the poor: Child poverty rises as the safety net shrinks.*
Julie Strawn.

No. 4. (Winter) *Crack's children: The consequences of maternal cocaine abuse.*
Theresa Lawton Hawley and Elizabeth Disney.

Volume VII (1993)

No. 1. (Spring) *Canadian special education policies: Children with learning disabilities in a bilingual and multicultural society.*
Linda S. Siegel and Judith Wiener.

No. 3. (Fall) *Child witnesses: Translating research into policy.*
Stephen J. Ceci and Maggie Bruck.

No. 4. [Winter, tentative title]. Integrating science and ethics in research with high-risk and underserved children and youth.
Celia B. Fisher.

Subscriptions ($10 to nonmembers of SRCD) and single issues ($3.50) are available from:

SRCD Executive Office
University of Michigan
300 North Ingalls, 10th floor
Ann Arbor, MI 48109-0406

Social Policy Report *is a quarterly publication of the Committee on Child Development, Public Policy and Public Information of the Society for Research in Child Development. Copyright of the articles published in the* Report *is maintained by the originator[s] of the work. Statements appearing in the* Report *are the views of the author and do not imply endorsement by the Editor or by SRCD.*

Editor: Nancy G. Thomas
Phone: (303) 925-5516
Fax: (303) 925-9570

SRCD Executive Office
University of Michigan
300 N. Ingalls 10th Floor
Ann Arbor, MI 48109-0406

SOCIAL POLICY REPORT

SOCIAL POLICY REPORT

Society For Research In Child Development

Volume VII, Number 3 Fall 1993

CHILD WITNESSES: TRANSLATING RESEARCH INTO POLICY

Stephen J. Ceci
Maggie Bruck

On August 12, 1983, Judy Johnson, the mother of a toddler at a prestigious nursery school in Manhattan Beach, California, told police that her 2-year-old son had been molested by Raymond Buckey, a teacher and the grandson of the school's founder, Virginia McMartin. Buckey was arrested but subsequently released for lack of evidence. On March 22, 1984, he was indicted by a grand jury and rearrested along with six female teachers, including his mother, Peggy McMartin Buckey. He was held without bail until 1989.

In August, 1984, the first of many preliminary hearings had begun for the seven McMartin Preschool defendants. Fourteen former students at the nursery school took the witness stand at this hearing and described a series of bizarre events involving sexual abuse, satanic rituals, and animal mutilation that allegedly occurred at their preschool. Based on these children's testimony, the defendants were accused of 115 counts of abuse, later expanded to 321 counts, including rape, sodomy, fondling, oral copulation , and drugging, and photographing of at least 100 children in the nude.

In January 1986, after 17 months of preliminary hearings in which each child witness was cross-examined by each of seven different defense attorneys, charges against five of the six women were dropped because of insufficient evidence. Only 26-year-old Raymond Buckey and his 58-year-old mother, Peggy McMartin Buckey, remained as defendants. After spending 2 years in jail, Peggy McMartin Buckey's bail was set at $500,000 in 1986, while her son remained remanded without bail in a proceeding that would last an additional 4 years.

In the ensuing years of legal proceedings, the major issue before the court was whether or not to believe the children. On the one hand, it was argued that the children's reports were authentic and that their bizarre and chilling accounts of events, which were well beyond the realm of most preschoolers' knowledge and experience, only served to substantiate the fact that the children had

PS 024022

actually participated in them. On the other hand, it was argued that the children's reports were the product of repeated suggestive interviews by parents, law enforcement officials, and therapists, and that the children were only reporting events suggested to them during these interviews.

In January 1990, following a 33-month trial and 9 weeks of deliberation, the jury in the McMartin Preschool sexual abuse case returned "not guilty" verdicts on 52 of the 65 counts. The jury deadlocked on 12 molestation charges against Raymond Buckey and on 1 count of conspiracy against his mother. Judge William Pounders dismissed the conspiracy charge.

In response to the acquittals, the children's parents railed at the way the case had been handled and at the jury's verdicts. The children themselves appeared on nationally televised talk shows, weeping over the jury's seeming refusal to believe their claims of ritualistic abuse. Newspapers and magazines across the nation ran headlines such as: "Doubt the children and jail the parents!"

During a postverdict press conference, many jurors claimed that they believed that some of the children had been abused, but were unable to reach a guilty verdict because of the suggestive way the children had been interviewed. These jurors claimed that the social workers prevented the children from speaking in their own language and thus diminished their credibility.[1]

By January 1990, the prosecution announced Buckey would be retried on the remaining 12 charges on which the jurors had been deadlocked. Although this trial was shorter, the jury again deadlocked. When the judge declared a mistrial, the prosecution did not to retry the case.

The McMartin case made legal history. In sheer magnitude, it was without parallel, lasting 7 years from the time of its inception to the final verdicts, producing hundreds of thousands of pages of transcript, and costing the State of California over $16 million. From the very start, the McMartin case captured the attention of the national media, with regular accounts of the children's allegations appearing on television (e.g., ABC News *Nightline*, April 20, 1984), in newspapers (e.g., Charlier & Downing, 1988; Shaw, 1990) and in magazine articles (e.g., Fischer, 1988).

The McMartin case is not a singular happening. There have been many similar cases in North America and Europe, some of which have received extensive media attention.[2] For the most part, these cases share the following elements: First, the witnesses were preschoolers at the time of the alleged abuse. Second, the disclosures were not made immediately following the alleged event, but after a long delay. Third, the disclosures often were preceded by intensive interviewing of the children by various professionals (e.g., child protective service workers, law enforcement) and nonprofessionals (e.g., parents, grandparents). Fourth, the children were the only witnesses to these alleged events, and corroborative physical evidence was lacking. Fifth, none of the defendants ever made a confession; all maintained their innocence, even after some co-defendants were convicted. Finally, the major issue before the jury in all of these cases was whether to believe the children.

Researchers in child development have served as witnesses or consultants for the defense and/or the prosecution in all of these trials. More importantly, these cases have changed the course of our research and thinking on children's memory development. They have encouraged researchers to tackle new issues, to develop innovative experimental paradigms, and to challenge and elaborate previous research on the reliability of young children's statements.

In this *Social Policy Report* we provide a glimpse into the social science research that has accumulated on the aspect of children's testimony that figured so prominently in the McMartin case and hundreds of others like it,[3] namely, preschoolers' presumed suggestibility. We then present some tentative thoughts about the policy implications of this research, by addressing the questions: What can (and should) researchers and mental health professionals tell courts when they are called upon to serve as expert witnesses and consultants? What is the proper role for professional organizations to play in overseeing expert testimony? First, however, we provide some background information about the problem of child sexual abuse in the U.S. and about the history of children providing courtroom testimony.

Prevalence of Abuse and Court Involvement

Crime statistics reflecting the sexual abuse of children are of great social concern. These statistics come from two major sources: one is based on annual rates of "substantiated" or "indicated" reports of child abuse,[4] and the other is based on adults' reports of abusive events during their own childhoods. The first source thus provides estimates of *incidence*, whereas the second provides estimates of *prevalence*.

According to the most recent *incidence* figures (based on data from 45 states), there were 2.7 million reports of suspected child maltreatment in 1991 (National Center for Child Abuse and Neglect, 1993); 129,697 of these were substantiated or indicated cases that were sexual in nature, indicating an incidence of childhood sexual abuse of less than 1% of children for that year. The accuracy of these rates has been challenged, (e.g., Besharov, 1991; Robin, 1991). On the one hand, these figures may overestimate the extent of child sexual abuse because they include indicated and substantiated cases which are not validated. Others, however, argue that these rates may underestimate the incidence of child sexual abuse, because many cases of actual child sexual abuse end up being classified "unsubstantiated," or significant numbers of cases of abuse are never reported to authorities. It seems plausible that while some substantiated cases are actually false, many more unsubstantiated and unreported cases are real. Thus, the national data likely underestimate the true incidence, although no one can say by how much.

Random samples of adults, asked about their childhood history, yield highly variable estimates of childhood sexual abuse, ranging, for females, from 6.8% (Siegel, Sorenson, Golding, Burnam, & Stein, 1987) to 62% (Wyatt, 1985); and for males, from 3% to 31% (see Peters, Wyatt, & Finkelhor, 1986). A number of methodological factors may account for these discrepant figures, but review of these is beyond the scope of this article. Our point is that even if the lowest prevalence rates are the most accurate, and even if the incidence of child sexual abuse in 1991 was 1%, this still represents a serious societal problem.

As a result of society's reaction to these dramatic figures, and particularly in reaction to the ineffective prosecution of child abuse cases, the legal system has been forced to change some of its rules concerning the admissibility and treatment of child witnesses. For example, until recently courts of law in all English-speaking countries were reluctant to accept the uncorroborated statements of child witnesses (Chadbourn, 1978). This reluctance was reflected in competency hearings, corroboration requirements, and cautionary instructions that some judges gave to juries concerning the risk of convictions based solely on the testimony of child witnesses (Andrews, 1964; Cohen, 1975). During the 1980s, however, all but a few states dropped their corroboration requirement for children in sexual abuse cases, a crime that by its nature lacks corroboration. Seventeen states now allow children to testify regardless of the nature of the crime, permitting the jury to determine how much weight to give to the child's testimony. With the continued adoption by states of Federal Rules of Evidence, the number of child witnesses is likely to expand.

As more and more children have been admitted as witnesses in the courtroom, legal and courtroom procedures have been modified. For example, some courts have instituted shield laws which permit a child witness to testify either behind a one-way screen or over closed-circuit television, to occlude the child's view of the defendant but not the defendant's view of the child. Hearsay exceptions are also allowed, whereby therapists, pediatricians, and others describe what children have said to them. These measures serve to assist child witnesses who otherwise might be "psychologically unavailable" to testify in open court (McGough, in press; Montoya, 1992, 1993).

In light of claims that many of these modifications challenge the constitutional rights of defendants (*Maryland* v. *Craig*, 1990; *Coy* v. *Iowa*, 1988), it is important to determine whether such procedures do, in fact, facilitate the accuracy of children's testimony (Montoya, 1993). This is particularly important in light of the fact that recent court decisions regarding the treatment of child witnesses have not been predicated exclusively on humane issues (i.e., reducing the stress placed on child witnesses), but also on the presumption that courtroom modifications will increase testimonial accuracy (Harvard Law Review Notes, 1985; Montoya, 1992). No scientific data addressing these issues are yet available, although some data on the costs and benefits of courtroom innovations on children's courtroom behavior have been gathered (Batterman-Faunce & Goodman, 1993; Flin, 1993).

There are, however, no reliable national data on the impact of these changes on the number of children who actually end up participating in family court or criminal justice proceedings. Gray's recent analysis (1993) of eight jurisdictions suggests that 3% to 10% of all cases of sexual abuse that are eventually filed with police result in a trial. Based on the National Center on Child Abuse and Neglect (1993) statistics cited above, this would suggest that, in the 45 states reporting, up to 13,000 children testified in sexual abuse trials that year. For various reasons, however, the actual incidence of court involvement may be considerably higher. The majority of cases that end in pleas still require the child to be deposed even if the child does not testify in a court trial. Second, these figures do not include data from five states, one of which is New York where as many as 3,150 children had formal court involvement in sexual abuse cases in 1990 (Doris, 1993). Finally, if nonsexual types of abuse and nonabuse cases involving children serving as witnesses are included (e.g., cases of domestic violence, custody disputes, accidents, playground injuries), then the estimate of children's participation in the legal system rises considerably, possibly to over 100,000 cases annually. We must emphasize, however, that this is, at best, an educated guess.

One final point about children's court involvement is in order. It appears that preschoolers are disproportionately more likely to be abused and more likely to have their cases come to trial. In an analysis of a sample of nearly 800 alleged victims of child sexual abuse in New York, preschoolers (ages 6 and younger) accounted for nearly 40% of the official sexual abuse cases, and 28% were aged 5 and younger (Doris, 1993). In Gray's recent analysis (1993), children below the age of 8 accounted for 45% of sexual abuse cases, and 18% were 5 years old or younger; 31% of the cases involving 5- to 6-year-old children went to trial, and 24% of cases involving 3- to 4-year-olds went to trial, whereas only 10% of the cases involving 13- to 14-year-olds went to trial.

Despite modifications in the judicial system resulting in the greater court involvement of children, both jurists and social scientists continue to raise fundamental questions about whether these changes actually facilitate the accuracy of children's testimony (Montoya, 1992; 1993) and, more broadly, whether the testimony that children do give is accurate. Next, we discuss the research that has been carried out on one important aspect of the accuracy of children's reports: the degree to which very young children are disproportionately prone to suggestion.

Research on Children's Suggestibility: Past and Present Trends

The scientific research on the suggestibility of children's recollections is both contradictory and confusing. A review of 20th-century studies of children's suggestibility can be found in Ceci & Bruck, 1993. Our purpose in this report is not to recap that analysis but to highlight some of the salient conclusions of these studies and to focus on the different experimental approaches used in investigating children's suggestibility.

Early Studies of Suggestibility

Early studies of children's suggestibility, with few exceptions, led to a jaundiced portrayal of children's proneness to suggestibility. Beginning with the early experiments of Binet and his European colleagues (Binet, 1900; Lipmann, 1911; Stern, 1910; Varendonck, 1911), and concluding with empirical studies in the 1920s and 1930s (Messerschmidt, 1933; Otis, 1924; Sherman, 1925), early researchers viewed children as extremely susceptible to leading questions and unable to resist an interviewer's suggestions. M. R. Brown (1926), a legal scholar, wrote:

> Create, if you will, an idea of what the child is to hear or see, and the child is very likely to see or hear what you desire. (p. 133)

Although the conclusions of these early researchers were confirmed by studies conducted right up until the 1980s, modern researchers have been ambivalent about generalizing these results to the forensic arena for several reasons. First, despite the fact that there is great concern currently about the reliability of preschoolers' reports, not one study in the first 80 years of this century included preschoolers. More recent research has begun to fill this void; since 1980, over 20 studies relevant to the issues of children's suggestibility have included a preschool sample.

A second and more important concern was that most of the previous studies involved children's recall of events that were forensically irrelevant. In most of this earlier literature, researchers examined the influences of a single misleading suggestion or a leading question on children's reports of neutral, nonscripted, and often uninteresting events that occurred in a laboratory setting. Although these results may be of importance for theoretical conceptualizations of the mechanisms that underlie suggestibility effects and memory processes, they have limited practical and legal relevance to the reliability of the child witness. In many court cases, the allegations involve the child as participant and not as bystander; they involve the child's recall of salient, rather than peripheral, events; they often involve repeated interviews which are highly suggestive; and they frequently involve emotionally charged and highly stressful events, such as sexual molestation. The earlier experiments of this century provide no clues as to the testimonial accuracy of children in such circumstances.

But how does a researcher conduct an ethically acceptable experiment that mirrors the many conditions that are characteristic of the child victim–witness? It would be unacceptable, for example, to determine whether an interviewer can successfully suggest to children with substantiated histories of abuse that the abuse had never taken place. Similarly, it would be unacceptable to determine if nonabused children will make allegations of sexual abuse after a highly suggestive interview. It is ethically impermissible to alter such fundamental aspects of young children's autobiography.

Modern Studies of Children's Suggestibility

In the past several years, a number of researchers have attempted to deal with these issues by developing new paradigms which admittedly do not mirror all of the conditions that bring children to court, but which do contain some important elements of the child witness's experiences. This section describes three major lines of recent research, each of which illustrates a different paradigm: (1) increasing the salience of the experienced events about which children will be interviewed, (2) increasing the dynamics of the interview situation, and (3) adding anatomically correct dolls to the interviewing context.

Increasing the salience of events. As discussed above, earlier studies were criticized as not forensically relevant because they did not examine how children respond to questions about events that involved their own body, or about other salient events that occurred in personally experienced and stressful situations. In response, a number of researchers have designed studies in which children are asked misleading questions about being touched. In some studies, children are questioned about their previous interactions with an experimenter in a laboratory (e.g., Rudy & Goodman, 1991). In other studies, children are questioned about an inoculation (Goodman, Hirschman, Hepps, & Rudy, 1991) or a genital examination (Saywitz, Goodman, Nicholas, & Moan, 1991).

For example, Saywitz and her colleagues (Saywitz et al., 1991) examined 5- and 7-year-old girls' memories of a medical examination. Half of each age group had a scoliosis exam (for curvature of the spine), and the other children had a genital exam. When children were interviewed either 1 or 4 weeks later, they were asked suggestive and nonsuggestive questions that were abuse-related (e.g., "How many times did the doctor kiss you?") or nonabuse-related (e.g., "Didn't the doctor look at your feet first?"). Although the older children were initially more accurate than the younger children on most questions, some of these age differences disappeared after the 4-week delay. Most importantly, although there were age differences in response to the suggestive abuse questions, very few children of either age gave incorrect responses; the 7-year-old children never made a false report of abuse, and the 5-year-olds did so only 4 times, although they were given 215 opportunities.

Saywitz and her colleagues point out specific patterns of results in this study. They conclude the children's inaccurate reports involved mainly errors of omission rather than commission. The majority of children in the genital examination condition did not disclose genital contact unless specifically asked to do so. This latter opportunity was only provided with the direct (leading) question format ("Did the doctor touch you here?"). In the scoliosis condition, when children were asked these direct questions, 2.86% of the children falsely affirmed vaginal touch and 5.56% falsely affirmed anal touch.[5] In reviewing this study, Goodman and Clarke-Stewart (1991) conclude that:

> . . . obtaining accurate testimony about sexual abuse from young children is a complex task. Part of the complexity rests in the fact that

> there are dangers as well as benefits in the use of leading questions with children. The benefits appear in the finding . . . that leading questions were often necessary to elicit information from children about actual events they had experienced (genital touching). . . . The children. . . . were generally accurate in reporting specific and personal things that had happened to them. If these results can be generalized to investigations of abuse, they suggest that normal children are unlikely to make up details of sexual acts when nothing abusive happened. They suggest that children will not easily yield to an interviewer's suggestion that something sexual occurred when in fact it did not, especially if nonintimidating interviewers ask questions children can comprehend. (pp. 102-103)

Thus, according to this group of researchers, earlier studies of children's suggestibility may have overestimated the extent to which they are suggestible. For example:

> There is now no real question that the law and many developmentalists were wrong in their assumption that children are highly vulnerable to suggestion, at least in regard to salient details. Although some developmentalists may be challenged to find developmental differences in suggestibility in increasingly arcane circumstances, as a practical matter who really cares whether 3-year-old children are less suggestible about peripheral details in events that they witnessed than are 4-year-old children? Perhaps the question has some significance for developmental theory, but surely it has little or no meaning for policy and practice in child protection and law. (Melton, 1992, p. 154)

It is important, however, to point out that not all data on children's reports of medical procedures are consistent with these conclusions. Ornstein and his colleagues (Baker-Ward, Gordon, Ornstein, Larus, & Clubb, in press; Ornstein, Gordon, & Larus, 1992) found that when children were later questioned about their memories of the visit to the pediatrician, 3-year-olds were more prone than 6-year-olds to make false claims in response to suggestive questions about silly events involving body contact (e.g., "Did the nurse lick your knee?"). Oates and Shrimpton (1991) also found that preschoolers were more suggestible than older children about previously experienced events that involved body touching. In contrast to the Saywitz et al. findings that false reports in response to suggestive questions are relatively infrequent, the younger children in these latter studies provided a substantial number of false reports in response to suggestive questions. Until recently, however, only a few studies have included explicit questions about sexual touching. Recent research by us and our colleagues has yielded different results, which will be reported in greater detail later when we describe our study of a pediatric examination.

Increasing the dynamics of the interview. A second major innovative theme in the current research on children's suggestibility involves examining the effects of various interviewing techniques on children's reports. This focus has arisen in response to the concern that the interviewing procedures of earlier studies were less intense than those that bring children to court—so much so as to result in a potential *underestimation* of children's suggestibility (Raskin & Esplin, 1991; Steller, 1991).

The interviewing procedures used in traditional laboratory studies and those used in the forensic arena differ in several ways. First, it is frequently the case that children who come to court are questioned weeks, months, or even years after the occurrence of an event, as opposed to minutes or days later. Suggestibility effects may be more salient after long delays, because the original memory trace has faded sufficiently to allow the suggestion to intrude more readily than might occur after shorter delays.

Second, child witnesses are rarely interviewed only one time, by one interviewer, or under nonstressful conditions. The modal child witness has been interviewed between 4 and 11 times prior to the first courtroom appearance; sometimes children are interviewed weekly for years about the same event—in therapy sessions, for instance. Leichtman and Ceci (1993) have suggested that the incessant use of leading questions and suggestions in these repeated interviews may result in a qualitatively different type of report distortion than that which arises from a single misleading question in a single postevent interview.

Third, an examination of the interviews of some child witnesses reveals that the label "suggestive interview" may describe more than the use of misleading questions. Rather, implicit and explicit suggestions can be woven into the fabric of the interview through the use of bribes, threats, repetitions of certain questions, and the induction of stereotypes and expectancies (Ceci & Bruck, 1993).

Finally, the questioning of child witnesses is typically conducted by parents, therapists, and legal officials, all of whom represent status and power in the eyes of the child; children may thus be more likely to comply with the suggestions of these interviewers than with those of the neutral interviewers employed in most research studies.

Although it is very difficult to create experimental conditions that simulate the confluence of the conditions present in child witness interviews (stressful episodes, with repeated and suggestive questioning over prolonged periods of time), researchers are beginning to examine how children's reports are influenced by the repetition of suggestions in multiple interviews prior to and following the occurrence of an event. In addition, researchers have focused on the interviewer and the potential effects that a particular interviewer's bias may have on the reports elicited from young children. We confine our discussion here to three studies recently carried out with our colleagues at Cornell and McGill universities, as they were designed specifically to address these issues (see Ceci & Bruck [1993], for discussion of additional studies). In focusing on our own studies, we necessarily present a particularized view, inspired by our own hypotheses, assumptions, and values, but the studies are designed to build on and challenge extant research.

In these studies, we patterned our experimental manipulations after materials collected over the past decade from court transcripts and from therapy sessions and law enforcement interviews involving children in cases similar to the McMartin case where there was a strong suspicion of abuse (see transcripts in Ceci, in press; Ceci & Bruck, 1993). These materials reveal that a child's first "disclosure" about abuse commonly occurs when an interviewer pursues a single hypothesis about the basis of the child's difficulties, which entails leading and suggestive interviews, often with fantasy inductions and "self-empowerment" techniques—the techniques themselves being potentially suggestive and stereotype inducing. Such disclosures are then pursued in law enforcement, child protective service, or therapeutic interviews.

Study 1: The effect of interviewer bias on children's reports

Ideally a forensic interview should be guided by a hypothesis-testing framework. Just as scientists try to arrive at the truth by ruling out rival hypotheses or by falsifying a favored hypothesis (Ceci & Bronfenbrenner, 1991; Dawes, 1992; Popper, 1962), interviewers should, in similar manner, attempt to rule out rival hypotheses, rather than exclusively attempting to confirm their favored one. However, because of situational pressures (e.g., case workers must sometimes make immediate determinations of potential danger to a child), it is not feasible that interviewers generate and test every conceivable hypothesis or, conversely, that they be "blind" to obviously relevant information pertaining to a main hypothesis that abuse is indeed present. Failure to recognize relevant information provided by the child could result in crucial missed opportunities.[6] But, as the following study shows, failure to test a rival hypothesis can result in reporting errors.

In this study (Ceci, Leichtman, & White, in press), we examined how an interviewer's hypothesis can influence the accuracy of young children's reports. Preschoolers were exposed to a game-like event and then interviewed 1 month later. The interviewer was given some information about events that *might* have occurred; some of the information was accurate and some of it was inaccurate. The interviewer was told to interview each child and to use whatever strategies she felt necessary to elicit the most factually accurate report from the child. The information we provided influenced the interviewer's hypotheses about what had transpired in this game, which, in turn, appeared to exercise a powerful influence on the dynamics of the interview, with the interviewer eventually shaping some of the children's reports to be consistent with her hypothesis about what had happened. When the interviewer was accurately informed, she got children to recall correctly 93% of the events that had transpired. It is important to note that the children made no false accusations when the interviewer was correctly informed, that is, they only made "errors of omission." However, when the interviewer was misinformed, 34% of the 3- to 4-year-olds and 18% of the 5- to 6-year-olds corroborated one or more false events that the interviewer erroneously believed had transpired. Thus, in the misinformed condition, the children made "errors of commission." Finally, the children seemingly became more credible as the interview unfolded. Many children initially stated details inconsistently, or with reluctance or even denial, but as the interviewer persisted in asking about nonevents, some abandoned their hesitancy and denials.

Because the interviewers were trained professionals (one was an experienced social worker, the other a nursery school teacher), we feel that the types of interactions observed in this study may be similar to those that occur in interviews between young children and parents, teachers, and professionals who are not given explicit training in how to generate and test alternative hypotheses. Our review of the materials from some publicized cases, such as McMartin, reveals that professional interviewers often steadfastly stick with one line of inquiry even when children continue to deny that the questioned events ever occurred (for examples see Ceci, in press).

Study 2: The effects of stereotype induction and repeated suggestions on young children's reports

A stranger named Sam Stone paid a 2-minute visit to preschoolers (aged 3 to 6) in their day-care center (Leichtman & Ceci, in press). Following Sam Stone's visit, the children were asked for details about the visit on four different occasions over a 10-week period. On each occasion, the interviewer refrained from using suggestive questions; she simply encouraged children to describe Sam Stone's visit in as much detail as possible. One month later, the children were interviewed a fifth time by a new interviewer, who first elicited a free narrative about the visit. Then, using probes, she asked about two "nonevents" which involved Sam Stone doing something to a teddy bear and a book. In reality, he never touched either item.

When asked in the fifth interview, "Did Sam Stone do anything to a book or a teddy bear?" most children accurately replied, "No." Only 10% of the youngest (3- to 4-year-old) children's answers contained claims that Sam Stone did anything to a book or teddy bear. When asked if they actually *saw* him do anything to the book or teddy bear, as opposed to "thinking they saw him do something," or "hearing he did something," now only 5% of their answers contained claims that anything occurred. Finally, when these 5% were gently challenged ("You didn't really see him do anything to the book/the teddy bear, did you?") only 2.5% still insisted on the reality of the fictional event. None of the older (5- to 6-year-old) children reported that they had seen Sam Stone do either of the fictional actions.

Another group of preschoolers was presented with a stereotype of Sam Stone before he ever visited their school. We did this to mimic the sort of stereotypes that some child witnesses have acquired about actual defendants. (In actual cases, for example, some children have been told repeatedly that the defendant did "bad things.") Each week, beginning a month prior to the visit, the children in our study were told a new Sam Stone story in which he was depicted as very clumsy. For example:

> You'll never guess who visited me last night. [pause] That's right. Sam Stone! And guess what he did this time? He asked to borrow my Barbie and when he was carrying her down the stairs, he tripped and fell and broke her arm. That Sam Stone is always getting into accidents and breaking things!

Following Sam Stone's visit, these children were interviewed four times over a 10-week period. These four interviews contained erroneous suggestions (e.g., "When Sam Stone ripped that book, was he being silly or was he angry?"). At the fifth interview, these children were asked for a free narrative about Sam's visit and were then asked probing questions about the two nonevents.

In this last interview, 72% of the youngest preschoolers claimed that Sam Stone did one or both misdeeds, a figure that dropped to 44% when they were asked if they actually saw him do these things. Importantly, 21% continued to insist that they saw him do these things, even when gently challenged. The older preschoolers, though more accurate, included 11% of children who insisted they saw him do the misdeeds.[7]

Some researchers have opined that the presence of perceptual details in reports is one of the indicators of an actual memory, as opposed to a confabulated one (Raskin & Yuille, 1989; Schooler, Gerhard, & Loftus, 1986). In this study, however, the presence of perceptual details was no assurance that the report was accurate. In fact, children in the stereotype plus suggestion condition produced a surprising number of fabricated perceptual details to embellish their false accounts of nonevents (e.g., claiming that Sam Stone took the teddy bear into a bathroom and soaked it in hot water before smearing it with a crayon). The difference in the quality of reports obtained in this study compared to others in the suggestibility literature may reflect the conditions under which the reports were obtained. As mentioned earlier, in most past studies, children's erroneous reports were in response to a single misleading question, posed after a brief delay following the event in question. In contrast, in the present study, children's false reports were a product of repeated erroneous suggestions over a relatively long period of time, coupled with a stereotype that was consistent with these suggestions.

It is one thing to demonstrate that children can be induced to make errors and include perceptual details in their reports, but it is another matter to show that their faulty reports are convincing to others. To examine the believability of the children's reports, we showed videotapes of their final interview to approximately 1,000 researchers and clinicians who work on children's testimonial issues. These researchers and clinicians were told that all the children observed Sam Stone's visit to their day-care centers. They were asked to decide which of the events reported by the children actually transpired and then to rate the overall credibility of each child.

The majority of the professionals were inaccurate. Analyses indicated that these experts—who conduct research on the credibility of children's reports, provide therapy to children suspected of having been abused, or carry out law enforcement interviews with children—generally failed to detect which of the children's claims were accurate, despite being confident in their judgments. Since so many of the children claimed that Sam Stone ripped the book and/or soiled the bear, it is understandable that many of the experts reasoned that these events must have transpired. But their overall credibility ratings of individual children were also highly inaccurate, with the very children who were least accurate being rated as most accurate. We believe that the highly credible yet inaccurate reports obtained from the children resulted from a combination of repeated interviews with persistent and intense suggestions that built on a set of prior expectations (i.e., a stereotype). In a similar way, it may become difficult to separate credibility from accuracy when children, after repeated interviews, give a formal videotaped interview or testify in court.

Study 3: Influencing children's reports of a pediatric visit

It could be argued that the Sam Stone Study is not relevant to evaluating the reliability of a child witness who reports personally experienced events involving his or her own body, especially when the experience involves some degree of distress. Furthermore, some might argue that the Sam Stone data are not germane to testimony about highly predictable and scripted events. In cases where the event involves a child's own body, is somewhat stressful, and is predictable, it is often thought that children may be less prone to suggestion.

To determine if children could be misled under such circumstances, we examined the influence of postevent suggestions on children's reports about a pediatric visit where they were examined (Ceci, Leichtman, & Bruck, in press). The study had two phases. In the first phase, 5-year-old children visited their pediatrician for an annual check-up. A male pediatrician examined the child. Then the child met a female research assistant who talked about a poster that was hanging on the wall in the examining room. Next, the pediatrician gave the child an oral polio vaccine and a DPT inoculation. Then the research assistant gave the child one of three types of feedback about how the child had acted when receiving the inoculation. One group was given pain-affirming feedback; they were told that it seemed as though the shot really hurt them, but shots hurt even big kids (hurt condition). A second group was given pain-denying information; these children were told that they acted like the shot did not hurt much, and that they were really brave (no-hurt condition). Finally, a third group was merely told that the shot was over (neutral condition). After the feedback, the research assistant gave each child a treat and then read the child a story. One week later, a different assistant visited the children and asked each one to indicate through the use of various rating scales how much he or she had cried during the shot and how much the shot hurt.

The children's reports did not differ as a function of feedback condition. Thus, we found that children could not be influenced to make inaccurate reports concerning significant and stressful procedures involving their own bodies. These results are similar in spirit to those of Saywitz et al. (1991) who also provided children with suggestions about stressful, personally experienced events in a single interview and discovered that children can be quite resistant to erroneous suggestions.

In the second phase of our study, we reinterviewed the children three more times, approximately 1 year after the shot. During these interviews, children were provided with repeated suggestions about how they had acted when they received their inoculations. Thus, as in the first phase of the study, some children were told that they were brave when they got their shot, whereas other children were not given any feedback. (For ethical reasons, we provided only "no-hurt" and "neutral" feedback in this phase of the study. We felt that providing "hurt' feedback might induce false or unpleasant memories about visiting the doctor.) When the children were visited for a fourth time and asked to rate how much the shot had hurt and how much they had cried, there were large suggestibility effects. Those who had been repeatedly told that they had acted brave when they had received their inoculation a year earlier reported significantly less crying and less hurt than children who were given no feedback. Thus, these data indicate that children's reports of stressful events involving their own bodies can be distorted under certain circumstances.

In the second phase of this study, we also tried to mislead children about the people who performed various actions during the original inoculation visit. Some children were falsely reminded on three occasions that the pediatrician gave them treats, showed them the poster, and read them a story. Some children were falsely reminded on three occasions that the research assistant gave them the inoculation and the oral vaccine. Control children were merely reminded that "someone" did these things. Based on the conclusions of other researchers (e.g., Fivush, 1993; Melton, 1992), it was hypothesized that children should not be suggestible about such important events and that they should be particularly immune to suggestions that incorporate shifts of gender. The male pediatrician had never given them treats or read them a story, and the female research assistant had never performed any medical procedures.

Contrary to these predictions, the children were misled. In the fourth interview, when asked about their doctor's visit in the previous year, 67% of the children (versus 27% of the control children) who were given misleading information about the pediatrician reported that the pediatrician showed them the poster, gave them treats, or read them a story. For children who were falsely told that the research assistant had given them the shot and the vaccine, 50% (versus 16% of the control children) fell sway to at least one of these two suggestions. Interestingly, 38% of the children who were given misleading information that the research assistant gave them the oral vaccine and the inoculation also said that the research assistant had performed other scripted events that not only had never occurred but also had never been suggested (e.g., reporting that the research assistant checked their ears and nose). None of the control children made such inaccurate reports. Thus, our suggestions influenced not only children's reports of personally experienced, salient events, but also their reports for nonsuggested scripted events that were related to the suggested events.

These data indicate that under certain circumstances children's reports concerning stressful events involving their own bodies can be influenced. The two factors that were most critical to this pattern of results were the intensity of the suggestions (i.e., repeating the suggestions over multiple interviews) and their timing (i.e., the long delay between the original event and interview about the event). These same two factors are characteristic of the conditions under which children made allegations of sexual abuse in many of the cases described at the beginning of this report.

The results of this study are consistent with the Sam Stone study even though the nature of the events about which children were misled were different. In the Sam Stone study, repeated suggestions and stereotypes led to convincing fabrications of nonoccurring events. In the pediatrician study, misleading information given in repeated interviews after a long delay following a target event influenced children's reports of personally-experienced, salient events.

The suggestibility of anatomically correct dolls. Anatomically correct dolls are frequently used by professionals, including child therapists, police, child protection workers, and attorneys, in interviewing children about suspected sexual abuse. According to recent surveys, 90% of field professionals use anatomical dolls at least occasionally in their investigative interviews with children suspected of having been sexually abused (Boat & Everson, 1988; Conte, Sorenson, Fogarty, & Rosa, 1991). Although no national figures are available, it appears that expert testimony is often based on observations of children's interactions with such dolls (Mason, 1991). We include a discussion here of anatomical dolls, because a number of commentators have raised questions about whether the dolls are suggestive (e.g., McGough, in press; Moss, 1988; Raskin & Yuille, 1989).

One rationale for the use of anatomical dolls is that they allow children to manipulate objects reminiscent of a sexual event, thereby cuing recall and overcoming language and memory problems. Another rationale is that their use is thought to overcome embarrassment and shyness. The dolls have also been used as projective tests. Some claim that if a child actively avoids these dolls, shows distress if they are undressed, or shows unusual preoccupation with their genitalia, this is consistent with the hypothesis that the child has been abused (see Mason, 1991).

The use of anatomically correct dolls has raised skepticism, however, among researchers and professionals alike. Two related arguments are frequently invoked against their use. The first is that the dolls are suggestive, that they encourage the child to engage in sexual play even if the child has not been sexually abused (e.g., Gardner, 1989; Terr, 1988). A child may insert a finger into a doll's genitalia, for instance, simply because of its novelty or "affordance," much the way a child may insert a finger into the hole in a doughnut. Another criticism is that it is impossible to make firm judgments about children's abuse status on the basis of their doll play because there are no normative data on nonabused children's doll play.

In several studies, researchers have compared the doll play of sexually abused and nonabused children. In addition, there have been a score of studies examining the doll play of nonabused children. Reviews of this literature (Berry & Skinner, 1993; Ceci & Bruck, 1993; Wolfner, Faust, & Dawes, 1993) indicate that many of the studies are methodologically inadequate and do not allow for firm interpretations about the potential usefulness or risks of using dolls. Furthermore, some data indicate that some of the play patterns thought to be characteristic of abused children, such as playing with the dolls in a suggestive or explicit sexual manner, or showing reticence or avoidance when presented with the dolls, also occur in samples of nonabused children (see Bruck & Ceci [1993] for a review). Finally, other data indicate that the dolls, though not suggestive, do not improve reporting—particularly among younger children (e.g., Goodman & Aman, 1990).

We have recently completed a study of 3-year-old children's interactions with anatomically correct dolls that highlights each of these results (Bruck, Ceci, Francœur, & Renick, in press). The children in this study visited their pediatrician for their annual check-up. The pediatrician conducted genital examinations with half the children; the remaining children did not receive genital exams. Immediately after the examination, the child was interviewed by a research assistant. Pointing to the buttocks and then to the genital areas of an anatomically correct doll, the assistant asked each child, "Did the doctor touch you here?" Later in the interview, the child was asked to use the doll to show how the doctor had touched his or her buttocks and genitals.

Children were quite inaccurate across all conditions. Only 45% of the children who received genital examinations correctly answered "Yes" to the questions "Did the doctor touch you here [on buttocks or genitals]?" Only 50% of the children who did not receive genital exams correctly replied "No" to these questions. Further, the children's accuracy did not improve when they were given the dolls and asked to show how the doctor had touched them. Only 25% of the children who had received genital examinations correctly showed how the pediatrician had touched their genitals and buttocks. (A significant number of female subjects in this condition were inaccurate, because they inserted their fingers into the anal or genital cavities of the dolls—which the pediatrician never did.) Only 45% of the children who did not receive genital examinations were accurate in not showing any touching; that is, 55% of the children who did not receive genital examinations falsely showed either genital or anal touching when given the dolls, a pattern most prevalent among the females in this group; 75% of the females who did not receive a genital examination falsely showed that the pediatrician touched their genitals or their buttocks.

With the data on the potential usefulness of dolls equivocal at best, we feel that an important confound in the literature deserves mention: the context for the presentation of the dolls in these research settings is very different from that of actual forensic and clinical settings. Transcripts of therapy sessions with children suspected of having been sexually abused reveal interviewers employing various practices: naming the dolls after defendants; berating the dolls for alleged abuses against the child (e.g., shaking a finger at the male doll who has been named after the defendant and yelling, "You are naughty for hurting Jennifer!"); assuming the role of fantasy characters in doll play; and creating a persistent atmosphere of accusation. In the research settings in which the use of anatomical dolls has been studied, nonabused children were never subjected to such highly suggestive experiences prior to being interviewed with the dolls; they were not given prior motivation to play with the dolls suggestively or aggressively. On the other hand, children who were alleged to have been abused were sometimes exposed to the dolls repeatedly prior to coming to the research setting; perhaps these interviews had involved repeated suggestions from parents and interviewers about various sexual themes. That their play with the dolls differed from that of nonabused children who lacked this prior experience could be attributed to the abused children's prior therapeutic or investigatory experiences, rather than to any inherent way in which they might be expected to play with the dolls.

Unfortunately, no study has examined the suggestive attributes of anatomical dolls, controlling for the preexperimental experience as a potentially serious confound. We simply do not know how nonabused children would behave with the dolls were they to have suggestive experiences prior to the experimental interview.[8] Conversely, we also do not know how abused children play with the dolls in their *first* investigatory interview, since the children in these studies have often been interviewed more than once and some have been exposed to the dolls at least once, prior to the experimental interview.

On the basis of our literature review (Ceci & Bruck, 1993), we concluded that the inconsistent findings point to the need for additional research and to the need for the development of explicit procedures to govern the use of anatomically correct dolls by interviewers. Until such research is available, the dolls ought to be used with great caution. Recently, Berry, & Skinner (1993) and Wolfner and his colleagues (1993) were even less supportive of doll use:

> . . . we are left with the conclusion that there is simply no scientific evidence available that would justify clinical or forensic diagnosis of abuse on the basis of the dolls. The common counter is that such play is "just one component" in reaching such a diagnosis based on a "full clinical" picture. . . . [Doll] play cannot be validly used as a component, however, unless it provides incremental validity, and there is virtually no evidence that it does." (Wolfner et al., p. 9)

Summary of current literature. The studies reviewed here highlight the different paradigms that researchers are now employing to examine children's suggestibility. In our review of this literature (Ceci & Bruck, 1993), we found that results of the most recent studies, in contrast to older ones, are somewhat more contradictory about the reliability of children's reports. One can locate studies claiming that young children are as immune to suggestion as older children (e.g., Marin, Holmes, Guth, & Kovac, 1979; Saywitz, et al., 1991), and studies claiming that younger children are more suggestible (Ceci, Ross, & Toglia, 1987; Cohen & Harnick, 1980; King & Yuille, 1987). Such mixed results have led to a confusing juxtaposition of headlines: "Study shows children are credible as witnesses." Or, "Research shows child witnesses unable to distinguish reality from fantasy."

A careful reading of the literature suggests, however, that there *are* reliable age differences in suggestibility, with preschoolers' reports more influenced by erroneous suggestions than older children's. In our review of the suggestibility literature, we found 18 studies that compared preschoolers to older children or to adults; in 15 of 18 of these studies, suggestibility was greater among preschoolers than older children or adults (see Table 2 in Ceci & Bruck, 1993). To be sure, some researchers attach various caveats to this conclusion. For example, some have claimed that age differences in suggestibility are evident mainly for nonparticipant children, i.e., bystanders (Rudy & Goodman, 1991); and for peripheral, nonsalient events (Fivush, 1993). And some researchers find that although young children may make some errors in response to suggestive questions with a sexual theme, on the whole they are highly resistant to such questions (e.g., Saywitz et al., 1991; Goodman et al., 1991). Still others have found larger age differences in suggestibility for questions with sexual themes (e.g., Baker-Ward et al., in press; Goodman, Rudy, Bottoms, & Aman, 1990) and for questions about salient events (e.g., Cassel & Bjorklund, 1993).

Although preschoolers are usually depicted as being the most suggestible, it is important to point out that older children and adults are also suggestible. For example, as described above, 7-year-olds' reports, after 1 year, of their visits to the pediatrician could be quite easily altered through suggestion. Clarke-Stewart, Thompson, & Lepore (reported in Goodman & Clarke-Stewart, 1991) also found that 7-year-old children's reports and interpretations of a recently experienced event could be easily manipulated through suggestion. Also, Goodman, Wilson, Hazan, & Reed (1989) found that a substantial number of 7- to 10-year-old children incorrectly agreed with interviewers' suggestions about details of an event that occurred 4 years earlier. Many of these misleading suggestions had sexual themes. Finally, suggestions can alter some fundamental aspects of adults' autobiographical memories (Loftus, 1993). Thus, we cannot conclude that older children and adults are not suggestible, only that their level of suggestibility is less than that of preschoolers.

We reiterate, however, that the conditions created in these studies differ markedly from those that occur in actual therapy or in law enforcement investigations: these latter two contexts are seldom as sanitized of affect and free of motives as those in the research setting. The real life situation may entail high levels of stress, assaults to the child's body, and loss of control. In some cases, children are interviewed and reinterviewed under emotionally charged circumstances, entailing the use of bribes and threats, and often in the presence of highly distressed parents; under such conditions some children may finally utter reports that are simply consistent with the interviewer's expectations. In the McMartin case, interviewers were alleged to have coerced children's statements by praising them when they reported events that were consistent with the interviewer's beliefs and criticizing them for failing to do so (e.g., calling them "dumb"). Interviewers in both this case and other day-care cases also told children that other children had already disclosed the details of the abuse, thus creating added pressure to assent to suggestions of abuse. Not surprisingly, interviewers in the McMartin case managed to elicit statements of abuse from 369 of nearly 400 children they interviewed (Sauer, 1993), although only one child had made claims of abuse prior to the interviews. (This girl's accusations were so bizarre that the prosecution dropped them from the case [Sauer, 1993]).

Elsewhere we and others have used more emotionally laden events to examine issues related to the role of affect and bodily touching in producing misinformation effects, including suggestions about being kissed while naked, witnessing parents violate norms, or hurting others to protect loved ones (see Ceci, Leichtman, Putnick, & Nightingale, 1993), and experiencing painful and/or embarrassing medical procedures (e.g., Goodman, 1993; Ornstein, Baker-Ward, Gordon, & Merritt, 1993). Although children's resistance to suggestions are sensitive to all of these factors (and others), no study has attempted to incorporate all of them into a single experiment.

It is highly unlikely, however, that we will ever mimic the assaultive nature of some acts or interviews perpetrated on child victims and witnesses. Thus we are far from being able to provide a definitive conclusion about the reliability of all child witnesses' reports. It is safe to conclude, though, that past pronouncements by some rather extreme advocates on both sides of the bench are simply unfounded. Children are neither as hypersuggestible and coachable as some prodefense advocates have alleged, nor as resistant to suggestions about their own bodies as some proprosecution advocates have claimed. They can be led, under certain conditions, to incorporate false suggestions into their accounts of even intimate bodily touching, but they can also be amazingly resistant to false suggestions and able to provide highly detailed and accurate reports of events that transpired weeks or months ago (e.g., Baker-Ward et al., in press). This mix of suggestibility and resistance to suggestion underscores the need for great caution in accepting the claims of those who would put either a prodefense or proprosecution "spin" on the data.

Policy Implications

Expert Witnesses

When a child comes to court to testify, this is often because he or she is the sole witness to a crime; this is particularly likely to be the situation in sexual abuse cases where the child is not only the sole witness, but there may be no physical evidence of abuse. The problem of uncorroborated testimony is compounded by the fact that the testimony of children may at times seem to lack credibility. As a result, both the prosecution and defense may call physicians, mental health professionals, and social scientists to serve as expert witnesses. In this section, we discuss the qualifications and roles of mental health professionals and social scientists who serve as expert witnesses in cases involving child witnesses, particularly in cases of alleged sexual abuse.

According to legal views (see Mason, 1991; Myers, 1993), these expert witnesses can be classified into two categories. (1) The first type, usually a mental health professional, is asked either to provide a generic description of the behavioral symptoms associated with sexual abuse, or to provide an opinion as to whether or not a particular child was abused. In the latter instance, the expert may have formulated his or her opinion based on therapy with the child or an assessment of a child's behavior; in some cases, the expert witness may have had no contact with the child in question. This type of expert is also frequently called upon to rehabilitate the credibility of a child witness who has been attacked by the defense, e.g., for delayed reporting. In this situation, the expert witness explains that, though such behaviors are *not* themselves diagnostic of abuse, it is not unusual for abused children to display a range of behaviors, such as recantation, delay of reporting, and inconsistent reporting. (2) The second type of expert witness is called to review the scientific literature on issues relevant to the credibility of child witnesses. This expert may cover various topics, including the literature on suggestibility as well as that on cognitive, emotional, and social development.

Although one might conclude that the research on children's suggestibility, discussed in this report, has policy implications for only the second type of witness, we argue that knowledge of this research is relevant to the professional qualifications and testimony of the first type of expert witness as well. The mental health professional who testifies on the diagnosis of sexual abuse or who describes to a court the symptoms associated with sexual abuse must also take into consideration competing hypotheses that might explain why the child in question, or children in general, demonstrate particular symptoms or make allegations of sexual abuse. One of the alternative hypotheses to be considered is that the particular child's allegations or symptoms have resulted from suggestive influences of the sort described above. It is important for the expert to consider such an alternative, because those same symptoms associated with sexual abuse (delayed reporting, retraction of the allegation, inconsistent accounts, inappropriate knowledge of sexual behavior, or unusual play with anatomically correct dolls) have been observed in nonabused children who have been exposed to suggestive influences (see Berliner & Conte, 1993; Kendall-Tackett, Williams, & Finkelhor, 1993, for recent reviews of the literature).

Diagnosing child sexual abuse is thus a complex task requiring experience with sexually abused children and knowledge of both the clinical and the suggestibility and developmental literature. Experts who testify on such matters should be well-versed in these domains (Myers, 1993).

Some legal scholars have pointed out, however, that little experience and training is required of expert witnesses who provide testimony to rehabilitate the child witness. Myers (1993) writes that because this type of testimony is thought straightforward and simple, "a child protection services worker with six months on the job and knowledge of three or four pertinent articles is qualified to provide rehabilitative testimony on recantation and delayed reporting" (p. 177). We argue that this witness should have a more thorough knowledge of the scientific literature on both the indicators of child sexual abuse and the literature on suggestibility.

The second type of expert witness, those who testify about the scientific literature on suggestibility and child development, does not require clinical experience. However, this expert must have a thorough knowledge of the research literature germane to his or her testimony.

What the expert witness on children's suggestibility should tell the court. We come now to the question that has vexed any social scientist who ever dreamed (or had nightmares) of being called upon to serve as an expert witness or to prepare an amicus brief for an appellate court on children's testimony, namely, what does our present state of scientific knowledge permit us to say about the reliability of the testimony of the child witness? Having acknowledged the complexities of the research, we hold that expert witnesses, regardless of whether they are testifying for the prosecution or for the defense, should cover several points based on the literature:

1. There are reliable age effects in children's suggestibility, with preschoolers being more vulnerable than older children to a variety of factors that contribute to unreliable reports.

2. Although young children are often accurate reporters, some do make mistakes—particularly when they undergo suggestive interviews; and these errors can relate not only to peripheral details, but also to salient, predictable events that involve their own bodies.

3. Measures can be taken to lessen the risk of suggestibility effects. To date, the factors that we know most about concern the nature of the interview itself—its frequency, degree of suggestiveness, and demand characteristics.

- A child's report is less likely to be distorted, for example, after one interview than after several interviews (the term "interviews" here includes informal conversations between parents and child about the target events).
- Interviewers who ask nonleading questions, who do not have a confirmatory bias (i.e., an attachment to a single hypothesis), and who do not repeat close-ended, yes/no questions within or across interviews, are more likely to obtain accurate reports from children.

- Interviewers who are patient, nonjudgmental, and who do not attempt to create demand characteristics (e.g., by providing subtle rewards for certain responses) are likely to elicit the best quality reports from young children.

Thus, at one extreme we can have more confidence in a child's spontaneous statements made prior to any attempt by an adult to elicit what they suspect may be the truth. At the other extreme, we are more likely to be concerned when a child has made a statement after prolonged, repeated, suggestive interviews. Unfortunately, most cases lie between these extremes and require a case-by-case analysis.

4. Finally, it is also important that the court appreciate the complexity of the interrelationships of the factors affecting children's suggestibility. As in most areas of social science, effects are rarely as straightforward as one might wish. Even though suggestibility effects may be robust, the effects are not universal. Results vary between studies, and children's behavior varies within studies. Thus, even in studies with pronounced suggestibility effects, there are always some children who are highly resistant to suggestion. Some studies may show reliable age differences in suggestibility even though the majority of both younger and older children did not succumb to suggestion. We have seen this in our own studies as well as in transcripts of forensic and therapeutic interviews: in some cases, no matter how much an interviewer may try to suggest that an event occurred, some children will consistently resist and not incorporate the interviewer's suggestion or point of view. On the other side, although suggestibility effects tend to be most dramatic after prolonged and repeated interviewing, some children incorporate suggestions quickly, even after one short interview (e.g., Clarke-Stewart, et al., 1989, as reported in Goodman & Clarke-Stewart, 1991). No facile conclusion can be presented to courts on this matter.

Ideal vs. Actual Expert Witnesses. The "model" expert witness who comes forward to testify on issues related to children's suggestibility should be someone who has thoroughly reviewed the pertinent literature and who can present the relevant facts in a balanced manner to the triers of fact. This requirement is not an easy one to meet; this research area is developing rapidly and is riddled with a host of complex issues that necessitate a broad understanding of design, statistics, and theory not likely possessed by someone outside the research community.

Unfortunately, many who serve as expert witnesses do not have this breadth of knowledge. We have reviewed many examples of testimony by so-called experts that appeared to have been based on incomplete and at times dubious knowledge. Nowhere in their testimony is there any hint of the complexities that ought to have tamed the witness's statements to the jury. In the worst cases, the testimony was actually opposite to what we know to be the best evidence from systematic research. All too often such an expert appears in court strictly because his or her opinion is consistent with that of the defense or the prosecution, rather than because the witness is truly knowledgeable about the field. Such testimony can be a disservice to the aims of justice, not to mention to the professions these expert witnesses represent.

Although the above discussion pertains mainly to the social scientist who testifies about children's suggestibility, our review of case material and the literature suggests that these same criticisms can be made of some experts who testify about the behavioral symptoms associated with sexual abuse. Mason (1991) analyzed 122 civil and criminal appellate court cases in which expert witnesses testified about child sexual abuse. She found that experts frequently presented testimony that was either internally inconsistent or was contradicted by other experts. For example, 14 experts cited age-inappropriate knowledge of sex and sexual preoccupations as characteristics of an abused child, whereas 6 experts asserted that naivete and aversion to sexual matters characterized the sexually abused child. Some experts maintained that consistent accounts of events were important indicators of sexual abuse, whereas others maintained the opposite, that sexually abused children are characterized by their inconsistent accounts. Mason also reported that appellate courts tend to take expert testimony at face value; that they rarely raise questions about the testimony's acceptance by the scientific community, or about the credentials of the mental health professional presenting expert testimony.

The response of the courts in this study reveals that there is a critical gulf between the scientific community and the judiciary. Judges are not willing and probably not able to critically evaluate the reliability of the testimony offered (Mason, 1991, p. 205).

An example of the unscientific nature of some experts' testimony is illustrated by *Kelly Michaels* v. *State of New Jersey*. Michaels was a preschool teacher convicted on 115 counts of sexual offenses involving 20 children, and sentenced to 47 years in prison. The expert witness for the prosecution testified that conduct of all but one of the child witnesses was consistent with having been sexually abused. She did not seriously consider the possible effects of numerous suggestive influences on the children's testimony, which had included persistent, aggressive, and suggestive interviews with children who initially denied that anything had happened. In this case, the appellate court did challenge the testimony of the expert witness, reversing Michael's conviction (after she had spent 5 years in prison), in part because the expert's testimony concerning the child behavioral indicators of abuse did not have acceptance within the relevant scientific community.

Problems with expert testimony are endemic to our legal system and to those of other countries whose codes have been derived from common law. As far back as one can check, jurists and laypersons alike have viewed expert witnesses as untrustworthy, as inclined to put a "spin" on interpretations of the data toward the side that hired them:

> Perhaps the testimony which least deserves credit with a jury is that of skilled witnesses. . . . It is often quite surprising to see with what facility, and to what extent, their views can be made to correspond with the wishes and interests of the parties who call them. (Judge John Pitt Taylor, 1858, p. 65-69, as quoted in Gross, 1991).

These views continue to be expressed by American jurists:

> To put it bluntly, in many professions service as an expert witness is not generally considered honest work. . . . Experts in fields see lawyers as unprincipled manipulators of their disciplines, and lawyers and experts alike see expert witnesses—those members of the learned professions who will consort with lawyers—as whores. (Gross, 1991, p. 1115)

And by British jurists:

> Expert evidence is sometimes given by people whose level of knowledge seems lamentably low. A number of the recent, and best-known scandals show this. . . . How does this come about? In the

> first place. I think it is because our present system provides no systematic quality control. Broadly speaking, anyone can be an expert witness, provided they have some relevant knowledge, and nothing whatever is done to see that only the best people are used. To be allowed to give expert evidence, witnesses must satisfy the judge that they have some practical experience, or some professional qualifications; but that is all. No minimum standards are laid down. The only test is opposing counsel's cross-examination; and, in a jury trial, this may be designed to score clever points, rather than to test whether they (i.e., the experts) are really good at their job. (Spencer, 1992, pp. 216-217)

The Relationship of Research to Clinical Practice

That the judicial community is unwilling or unable to evaluate critically the testimony of social science experts and mental health professionals reflects to some degree an incomplete or inaccurate understanding of the relevant knowledge base, but it also reflects, in the case of the reliability of children's reports, a gulf between clinical practice and social science research. As a result, in those cases where clinical practice is not informed by research findings, clinicians and social scientists may present diametrically opposite expert testimony on the very same topic. Two examples illustrate this breach between the two perspectives.

The first involves a survey of 212 mental health professionals about their assessment and validation procedures in sexual abuse cases (Conte et al., 1991). Of relevance to the present report, it was found that children had already been asked to tell their story an average of 2.3 times before talking to the professional respondent; only 27% of respondents indicated that they were the first person to talk with the child about the abuse. In discussing these findings, however, the authors do not seriously consider the impact of such interviewing practices:

> Little is currently known about the effects of such prior interviewing on the child's willingness to engage with yet another adult or on the quality of information obtained from the child. While some professionals are likely to make much of the possible "contamination" that these prior interviewers have on the child's reports, there are virtually no data currently available suggesting that adults have the power through interviewing techniques to alter fundamentally a child's understanding of and ability to describe what events did or did not take place. (p. 433).

We hope that this report will begin to inform professionals that such data are available.

A second example that illustrates the gulf between practice and research concerns the use of anatomically correct dolls. Many professionals have no formal training or experience in the use of the dolls (Boat & Everson, 1988) and may view some interactions of children with the dolls (e.g., placing a finger in the doll's anal cavity, tugging on its penis, or avoiding the dolls altogether) as indicative of sexual abuse, even though there is no scientific support that such interactions are diagnostic of abuse. In a recent survey, for example, only 16% of mental health and law professionals stated that avoidance of the dolls was normal, while 80% rated digital penetration as abnormal (Kendall-Tackett, 1991). Yet, as reviewed above, such behaviors are commonly observed in nonabused children.

Of more concern, perhaps, is the American Psychological Association's (1991) current position on the use of the dolls. The following statement was issued by APA's Council of Representatives:

> Neither the dolls nor their use are standardized or accompanied by normative data. . . . We urge continued research in quest of more and better data regarding the stimulus properties of such dolls and normative behavior of abused and nonabused children. . . . Nevertheless, doll-centered assessment of children, when used as part of a psychological evaluation and interpreted by experienced and competent examiners, may be the best available practical solution for a pressing and frequent clinical problem. (APA, 1991, p. 1).

The APA's policy position seems contradictory in its noting first that there are no standardized methods for doll interviews or normative data on nonabused or abused children's doll play, but then asserting that experienced interviewers may nevertheless find doll-centered assessment the best available method for evaluating children suspected of having been sexually abuse. Even if one assumes that experienced examiners can avoid making false inferences from children's doll play, and that such doll play can provide important clinical insights not obtainable from other sources, the APA should nevertheless codify this expert knowledge in such a way that researchers can accurately assess the incremental validity of doll-based assessments. Our reading of the literature is that at present such knowledge is more illusory than real (see Wolfner et al.'s criticism [1993] of the lack of incremental validity of doll-based assessments). Even if anatomical dolls are used as just one part of an assessment, other aspects of so-called "developmentally sensitive assessments" (e.g., play therapy, role playing, techniques that induce visually-guided imagery, self-empowerment training) may interact with the doll use to produce false positive assertions of abuse. Because the appropriate research has yet to be done, it is shortsighted to assume (as some experts have testified in court) that the dolls do not present reliability risks. Although it could be the case that the use of dolls does provide important information, it could also be the case that this method leads to unacceptable levels of false positive reports. Only research will tell.

The fact remains that clinicians and mental health professionals face many dilemmas and choices in providing for children who may have been sexually abused. Often the favored choice may conflict with forensic procedures. Let us consider one scenario: a child has been removed from her home as a result of a report of sexual abuse and has been placed in emergency foster care, separated from her family, friends, and school. The child is greatly distressed and in need of immediate counseling. The forensic interviews will not be completed for several months. In light of some research findings that children's reports are likely to be more accurate if interviews (which include therapy sessions) concerning the alleged abuse are held to a minimum until after the forensic interview takes place, when should the mental health professional begin therapy with the child? How can we avoid the twin dangers of, on the one hand, putting the child's emotional needs on hold until after the forensic interviews and, on the other hand, providing counseling that can be potentially damaging to the veracity of the child's report? We know of no easy answers.

Given the pressing needs of both sides in a criminal dispute to prepare, investigate, and often reinterview, no amount of child-friendly court procedures can totally alleviate some of the problems associated with children's testimony. Yet, perhaps there are ways of providing therapeutic support that lessen the likelihood of tainting the child's report. Therapeutic procedures that involve visually-guided imagery in the context of the abuse-related allegations might be avoided, as might forms of therapy that make contact with the abusive scenario (e.g., self-empowerment training, role playing, doll use, hypnosis).

Although some might argue that it would be too restrictive and ultimately damaging to a child's development were therapists to avoid potentially suggestive techniques, it could also be argued that employing such interventions simply constitutes too great a risk. On the one hand, if the defendant is innocent, such techniques could promote and reinforce false allegations. On the other hand, if the defendant is guilty, these interventions may end up discrediting the child's testimony, with defense attorneys arguing that the child's reports are the product of highly suggestive therapeutic techniques. Finally, on the empirical side, we are unaware of any persuasive treatment–outcome validity research indicating that suggestive techniques are necessary in therapy to achieve a positive mental health outcome for children suspected of being abused. Given this state of knowledge, clinicians might consider limiting interventions to *non*suggestive techniques in therapy until young clients have given sworn statements; such an approach may afford minimal danger to the child.

Professional Organizations as Ethical Gatekeepers

Professional organizations could help resolve some of the problems we have been discussing by making ethics codes for expert witnesses more explicit. Existing codes for expert witnesses of the organizations that represent various constituencies (psychology, social work, pediatrics) tend to be weak and ill-defined—in part because "expert witness" is an ill-defined legal concept.[9] The Federal Rules of Evidence 702 states that if scientific, technical, or other specialized knowledge will assist a fact finder in understanding evidence, then a witness may be regarded as an expert by virtue of his or her knowledge, skill, experience, training, or education. This rule construes expertise broadly enough to cover all fields, including emerging areas within fields, and is constrained by two other Federal Rules of Evidence (401 and 403), which specify that the expert testimony must be relevant. Together, these Federal Rules allow virtually anyone who possesses an advanced degree, or who has some clinical experience, to offer expert testimony on children's credibility, even though the expert may have scant knowledge of the current scientific findings. As a result, experts testifying in child sexual abuse cases have offered totally opposite interpretations of children's behavior and testimony (see Mason [1991] above).

To some extent, weak ethics codes also reflect the ascendancy of guild interests. Because no constituency wants to be excluded from activities that involve service to others (at times for financial gain), its representatives ensure that its members' role is not diminished by ethics code language. Ethics codes tend to be explicit about matters that are relatively benign to the group as a whole (e.g., rules for preparing reports, or statements regarding generic conflicts of interest), but vague about matters that could adversely affect the entire membership (e.g., defining precisely what an expert should know in order to testify about children's suggestibility, or what it means to conduct a good interview). Thus, for example, when psychologists look to their own specialty guidelines and general ethics codes for guidance about the credentials or conduct of an expert witness, they find little help other than enjoinders to act responsibly, to be informed, and to aspire to the norms that guide a professional toward the highest ideals. Consider some of the sections of the most recent APA code of ethics revision (Ethical Principles, 1992) relevant to forensic services:

- Psychologists appropriately take into account the ways in which a prior relationship might affect their professional objectivity or opinions and disclose potential conflict to the parties. (Section 705)
- Psychologists who engage in . . . professional activities maintain a reasonable level of awareness of scientific and professional information in their fields. (General Standards 1.5)
- Psychologists rely on scientifically and professionally derived knowledge when making scientific or professional judgments. (General Standards 1.6)
- In addition, psychologists base their forensic work on appropriate knowledge of a competence in the areas underlying such work. (General Standards 7.1)
- In forensic testimony and reports, psychologists . . . describe fairly the bases for their testimony and conclusions [and] whenever necessary to avoid misleading, acknowledge the limits of their data or conclusions. (General Standards 7.4).

However well intended, these statements taken together lend themselves to ambiguous interpretation. For example, can a psychologist, in testifying about children's suggestibility, rely on *either* research knowledge *or* clinical experience? Can a therapist be expected to avoid a conflict of interest and maintain sufficient objectivity to serve as an expert witness when he or she has had extended contact with the child? (Apparently so. Mason's analysis [1991] showed that many expert witnesses who testified in abuse cases were often the child's therapist, and only 13% of all experts had no prior relationship with the child.)

Missing from ethics codes and specialty guidelines for expert witnesses (e.g., Committee on Ethical Guidelines, 1991) is language that would specify that they bring to court more than an advanced degree, a supervised internship which had brought them into contact with sexually abused children, or other clinical experience whereby they had occasionally seen sexually abused clients in their practice. An expert testifying on children's suggestibility, and more generally on the credibility of child witnesses, should be intimately familiar with the systematic scholarship on the topic. Although it is not necessary for this expert to be a researcher, he or she needs to be at least a critical consumer of the research literature.

The failure of professional organizations to constitute and then to enforce principled guidelines has serious consequences. First, it can undermine the judicial system's confidence in the capacity of professionals to offer reliable testimony. More important, in criminal proceedings where the defendant faces incarceration, or in civil proceedings where the future placement of the child is at stake, the legitimacy of the expert's testimony can be critical to preserving the rights of both the child and the defendant.

To conclude, enforcement mechanisms are needed to ensure that expert testimony can be evaluated for its scientific merit. Until such mechanisms are openly advertised to all consumers of legal services, enjoinders to "stay informed" will probably do little to ebb the sorts of abuses reported by Mason (1991) and Spencer (1992). Because of their vagueness, professional ethics codes will be implemented more often in the breach than in the letter.

Conclusion

We have argued that the investigation of child sexual abuse allegations and expert testimony addressing such investigations are fraught with problems. Scientists have begun to contribute important insights to these problems, though clearly more research is needed. We have provided some troubling examples of how research has failed to inform practice, and how experts often go beyond what current scientific findings seem to warrant.

To be sure, those charged with investigating, reporting, and treating suspected child maltreatment face immense obstacles. These professionals are deeply aware of the pervasiveness of child sexual abuse, and the all too frequent ineffectiveness of prosecution. They know better than most the emotionally wrenching sequelae of abuse, especially intrafamilial abuse. And they are keenly aware that the evidence from research must always be tempered by real-world considerations, no matter what the "significance level" or "effect size" of a finding. Thus, many pressing, unresolved issues concerning the interviewing and treatment of individual children remain.

We presented a scenario highlighting the difficult task faced by many professionals on a daily basis—how to promote two goals that often conflict: how to provide for the child's mental health needs while simultaneously protecting the legal rights of the accused. While we must strive to uncover abuse, we must eschew interview processes that may promote false beliefs, fantasies, or fabrications—regardless of the nature of the initiating event. Just as we have argued above that it is unethical for social scientists to institute experimental manipulations that might change the fundamental nature of children's emotionally salient autobiographical memories, it is equally indefensible for therapists or forensic interviewers to cause such changes. The results of persistent erroneous suggestions and of failures to test alternative hypotheses can be lasting, as evidenced by the experiences and reactions of the child witnesses in the McMartin trials, described at the beginning of this report:

> No one who saw them will soon forget the frenzied faces of . . . former McMartin pupils (who) had spent their last six years—fully half their lives—instructed in the faith that they had been subjected, at ages 4 and 5, to unspeakable sexual horrors; this belief they had come to hold as the defining truth of their lives and identities. It is not surprising that these children should have wept and raved when the verdict was handed down denying all that they believed in. (Rabinowitz, 1990, p. 63)

Notes

[1] In view of the public outcry against the seeming refusal by jurors to believe the children, posttrial statements by these same jurors about believing some of the children's claims may have been self-serving. One close observer of the trial suggested this possibility to us.

[2] Several have been the focus of books. The Wee Care case involving Kelly Michaels was the source for several books, including *Naptime* (Manshel, 1990) and *Not My Child* (Crowley, 1990). Other cases have been detailed in television documentaries (e.g., the Little Rascals case involving Bob Kelly and five other defendants, which was the focus of three Frontline documentaries, e.g., *Loss of Innocence*); movies (e.g., the Country Walk day-care case in Miami, which was the basis of the movie *Unspeakable Acts*); and

magazine and newspaper articles (e.g., Nathan, 1987; Rabinowitz, 1990).

[3] Although this report focuses on the interviewing of alleged child sexual abuse victims, the literature reviewed is equally important to nonabuse cases that involve the child witness. Sexual abuse is of special interest, because this category of complaints appears to represent the single largest class of actions that eventuate in criminal court testimony (as opposed to neglect cases or custody disputes which are largely litigated in juvenile and family court systems). Our discussion centers on nursery school cases, because, although these cases represent only a small proportion of sexual abuse complaints, in absolute numbers they involve a large number of children (in the McMartin case, for instance, interviewers under contract to the State of California alleged the abuse of 369 children [Sauer, 1993]); moreover, day-care cases are relevant to the more general testimonial issues found in many nonday-care cases (i.e., repeated suggestive questioning, interviewer stereotypes, failure to test alternative hypotheses). Finally, because of their visibility, day-care cases are often more extensively documented.

[4] Cases are classified as *substantiated* or *indicated* based on how consistent the evidence from an investigation is with abuse; in most states this is a matter of caseworker judgment. States usually have a two-tiered system of classifying investigations as either substantiated/founded, on the one hand, or unsubstantiated/not founded, on the other. Some states use a third tier that is intermediate between substantiated and unsubstantiated, namely, "indicated." This term is given to cases in which the agency doing the investigation may have "reason to suspect" that abuse occurred, but the level of evidence does not rise to the level required for the designation "substantiated." The lowest level of evidence needed to substantiate a case is "some credible evidence," which is used by 18 states, while the highest level of evidence needed is "preponderance of evidence," which is used by 12 states. An intermediate level of evidence is used by an additional 12 states, and the remaining states use idiosyncratic terminology (see Figure 5 of National Center of Child Abuse and Neglect, 1993, p. 28). Thus, lower levels of evidence increase the possibility that, upon further investigation, a subsequent determination may be made that insufficient evidence exists to designate the presence of abuse. For these reasons it is important not to use the terms "substantiation," "indication," and "validation" interchangeably.

[5] Some have suggested that these two figures be summed to 8%. This assumes, however, that there were different children in the two categories, which is not clear from the published report. The breakdown reported here is that reported by Saywitz, et al. (1990).

[6] Courts have taken notice of the need to distinguish between an interviewer whose view reflects a strongly held expectation versus the interviewer who possesses relevant background information. For example, in *Idaho v. Wright* the Court accepted the argument contained in an amicus brief that "there is an important distinction between preconceptions that can cloud judgment, and background information that is needed for thorough evaluation of sexual abuse" (Amicus Brief to the Supreme Court in *Idaho v. Wright*, No. 89-260, p. 96).

[7] These data reveal an interesting disjunction with the reasoning that when children retract earlier claims of sexual abuse, this is indicative, if not diagnostic, of a truthful original report (Sgroi, 1982; Summit, 1983). In this study, it was often the case that children originally made false allegations, which they then, with gentle persuasion, recanted. Were this finding applicable to situations that are abuse-related—and we make no such claim here—it could be suggested that retraction might also be consistent with an erroneous original report.

[8] Pilot data from one subject addresses this question. A 3½-year-old nonabused girl was examined by a pediatrician. She was not given a genital examination. Immediately after the examination, when interviewed by the experimenter, she correctly said that the doctor had not touched her genitals or buttocks. Furthermore, when shown an anatomically correct doll and told to show how the doctor had touched her genitals and buttocks, she correctly stated that he had not touched her. Three days later, the same child was shown the anatomically correct doll and asked to show all the things that the doctor did to her in her previous visit. This time, she inserted a stick into the vagina of the doll. Upon further questioning, however, she said that the doctor did not do this. Three more days later, the child was asked to use the anatomically correct doll to show her father everything that had happened at the examination. This time, she hammered a stick into the doll's vagina and then inserted a toy earscope into the doll's anus. When asked if this really happened, she said "Yes it did." When her father and the experimenter both tried to debrief her with such statements as, "Your doctor doesn't do those things to little girls. You were just fooling. We know he didn't do those things," the subject clung tenaciously to her claims. Thus, for this one subject, repeated exposure to the doll, with minimal suggestions, resulted in highly sexualized play. It is critical that such a finding be replicated with a large, diverse sample to determine if this child's response is representative of nonabused children.

[9] For example, most courts disallow expert testimony that speaks directly to the *ultimate question*, that is, the defendant's

guilt or innocence. In some courts, however, expert witnesses are permitted to testify as to whether they believe the child was abused (see Myers, 1992). One would think that an expert's opinion that a particular child was abused might have the same effect as speaking to the child's credibility. This leads to confusion even among the legal scholars whom we have consulted, with one remarking that the courts' thinking regarding this issue is little more than "wordplay."

References

American Psychological Association. (1991). *Statement on the use of anatomically detailed dolls in forensic evaluations.* Washington, DC: APA Council of Representatives.

Andrews, J. A. (1964). The evidence of children. *Criminal Law Review, 64,* 769-777.

Baker-Ward, L., Gordon, B., Ornstein, P. A., Larus, D., & Clubb, P. (in press). Young children's long-term retention of a pediatric examination. *Child Development.*

Batterman-Faunce, J. M., & Goodman, G. S. (1993). Effects of context on the accuracy and suggestibility of child witnesses. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 301-330). New York: Guilford Press.

Berliner, L., & Conte, J. R. (1993). Sexual abuse evaluations: Conceptual and empirical obstacles. *Child Abuse and Neglect, 17,* 111-125.

Berry, K., & Skinner, L. G. (1993). Anatomically detailed dolls and the evaluations of child sexual abuse allegations: Psychometric considerations. *Law and Human Behavior, 17,* 399-422.

Besharov, D. (1991). Child abuse and neglect reporting and investigation: Policy guidelines for decision making. *Child and Youth Services, 15,* 35-50.

Binet, A. (1900). *La suggestibilité.* Paris: Schleicher Frères.

Boat, B., & Everson, M. (1988). The use of anatomical dolls among professionals in sexual abuse evaluations. *Child Abuse and Neglect, 12,* 171-186.

Brown, M. R. (1926). *Legal psychology.* Indianapolis: Bobbs-Merrill.

Bruck, M., Ceci, S. J., Francoeur, E., & Barr, R. (in press). *"I hardly cried when I got my shot": Young children's reports of their visit to a pediatrician.* Manuscript submitted for publication.

Cassel, W. S., & Bjorklund, D. F. (1993). *Tell me about . . . . Don't you remember . . . ? Isn't it true that . . . ? Developmental patterns of eyewitness responses to increasingly suggestive questions.* Unpublished manuscript, Florida Atlantic University, Boca Raton, FL.

Ceci, S. J. (in press). Cognitive and social factors in children's testimony. In B. Sales & G. VandenBos (eds.), *APA Master Lectures: Psychology and the Law.* Washington, DC.: American Psychological Association.

Ceci, S. J., & Bronfenbrenner, U. (1991). On the demise of everyday memory: The rumors of my death are greatly exaggerated. *American Psychologist, 46,* 27-31.

Ceci, S. J. & Bruck, M. (1993). The suggestibility of the child witness. *Psychological Bulletin, 113,* 403-439

Ceci, S. J., Leichtman, M., Putnick, M., & Nightingale, N. (1993). Age differences in suggestibility (pp. 117-138). In D. Cicchetti & S. Toth (Eds.), *Child witnesses, child abuse, and public policy.* Norwood, NJ: Ablex.

Ceci, S. J., Leichtman, M. D., & Bruck, M. (In press). The suggestibility of children's eyewitness reports: Methodological issues. In F. Weinert & W. Schneider (Eds.), *Memory development: State of the art and future directions* . Englewood Cliffs, NJ: Erlbaum.

Ceci, S. J., Leichtman, M., & White, T. (in press). Interviewing preschoolers: Remembrance of things planted. In D. P. Peters (Ed.), *The child witness in context: Cognitive, social, and legal perspectives.* Holland: Kluwer.

Ceci, S. J., Ross, D., & Toglia, M. (1987). Age differences in suggestibility: Psycholegal implications. *Journal of Experimental Psychology: General, 117,* 38-49.

Chadbourn, J. (1978). *Wigmore on evidence.* Boston: Little, Brown.

Charlier, T., & Downing, S. (1988, January 17-22). Justice abused: A 1980's witch hunt. Memphis *Commercial Appeal.*

Clarke-Stewart, A., Thompson, W., & Lepore, S. (1989, April). *Manipulating children's interpretations through interrogation.* Paper presented at the biennial meeting of the Society for Research in Child Development, Kansas City.

Cohen, R. L. (1975). Children's testimony and hearsay evidence. *Law Reform Commission of Canada. Report on evidence.* Ottawa: Information Canada.

Cohen, R. L., & Harnick, M. A. (1980). The susceptibility of child witnesses to suggestion. *Law and Human Behavior, 4,* 201-210.

Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law & Human Behavior, 15,* 655-665.

Conte, J. R., Sorenson, E., Fogarty, L., & Rosa, J. D. (1991). Evaluating children's reports of sexual abuse: Results from a survey of professionals. *Journal of Orthopsychiatry, 78,* 428-432.

Coy v. Iowa, 108 S. Ct. 2798 (1988).

Crowley, P. (1990). *Not my child: A mother confronts her* *child's sexual abuse.* New York: Doubleday.

Daro, D., & Mitchel, L. (1990). *Current trends in child abuse reporting and fatalities: The results of the 1989 annual 50 states survey.* Washington, DC: National Commission for the Prevention of Child Abuse.

Dawes, R. (1992, Spring). The importance of alternative hypothesis and hypothetical counterfactuals in general social science. *The General Psychologist,* 2-7.

Doris, J. (1993). Child witness conference. Supplemental RFP Children's Justice and Assistance Act funds. Submitted by the Family Life Development Center, Cornell University.

Ethical principles of psychologists and code of conduct. (1992). *American Psychologist, 47,* 1597-1611.

Federal Rules of Evidence, Rule 702.

Fischer, M. A. (1988). Flip-flop: Why four years later the press is taking a strikingly different approach to the McMartin preschool scandal. *Los Angeles Magazine, 33,* 85-94.

Fivush, R. (1993). Developmental perspectives on autobiographical recall. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 1-24). New York: Guilford Press.

Flin, R. (1993). Hearing and testing children's evidence. In G. S. Goodman & B. L. Bottoms (Eds.), *Child victims and child witnesses: Understanding and improving testimony* (pp. 279-300). New York: Guilford Press.

Gardner, R. (1989). *Sex abuse hysteria: Salem witch trials revisited.* Longwood, NJ: Creative Therapeutics Press.

Glaser, D., & Collins, C. (1989). The response of young, non-sexually abused children to anatomically correct dolls. *Journal of Child Psychology and Psychiatry, 30,* 547-560.

Goodman, G. S. (1993, March). *Children's memory for stressful events: Theoretical and developmental considerations.* Paper presented at the biennial meeting of the Society for Research in Child Development. New Orleans.

Goodman, G. S., & Aman, C. (1990). Children's use of anatomically detailed dolls to recount an event. *Child Development, 61,* 1859-1871.

Goodman, G. S. & Clarke-Stewart, A. (1991). Suggestibility in children's testimony: Implications for child sexual abuse investigations. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 92-105). Washington, DC: American Psychological Association.

Goodman, G. S., Hirschman, J. E., Hepps, D., & Rudy, L. (1991). Children's memory for stressful events. *Merrill Palmer Quarterly, 37,* 109-158.

Goodman, G. S., Rudy, L., Bottoms, B., & Aman, C. (1990). Children's concerns and memory: Issues of ecological validity in the study of children's eyewitness testimony. In R. Fivush & J. Hudson (Eds.), *Knowing and remembering in young children* (pp. 249-284). New York: Cambridge University Press.

Goodman, G. S., Wilson, M. E., Hazan, C., & Reed, R. S. (1989, April). *Children's testimony nearly four years after an event.* Paper presented at the annual meeting of the Eastern Psychological Association, Boston.

Gray, E. (1993). *Unequal justice: The prosecution of child sexual abuse.* New York: MacMillan.

Gross, S. R. (1991). Expert evidence. *Wisconsin Law Review, 1113,* 1114-1231.

Harvard Law Review Notes. (1985). The testimony of child sex abuse victims in sex abuse prosecutions: Two legislative innovations. *Harvard Law Review, 98,* 806-827.

Idaho v. *Wright,* 110 S. Ct. 3139 (1990).

Kendall-Tackett, K. (1991, April). *Professionals' standards of "normal" behavior with anatomical dolls and factors that influence these standards.* Paper presented at the biennial meeting of the Society for Research in Child Development, Seattle.

Kendall-Tackett, K. A., Williams, L. M., & Finkelhor, D. (1993). Impact of sexual abuse on children: A review and synthesis of recent empirical studies. *Psychological Bulletin, 113,* 164-180.

King, M., & Yuille, J. (1987). Suggestibility and the child witness. In S.J. Ceci, M. Toglia, & D. Ross, (Eds.), *Children's eyewitness memory* (pp. 24-35). New York: Springer-Verlag.

Leichtman, M. D., & Ceci, S. J. (revision under review). The effect of stereotypes and suggestions on preschoolers' reports. *Developmental Psychology.*

Lipmann, O. (1911). Pedagogical psychology of report. *Journal of Educational Psychology, 2,* 253-261.

Loftus, E. F. (1993). The reality of repressed memories. *American Psychologist, 48,* 518-537.

Manshel, L. (1990). *Nap time.* New York: Kensington.

Marin, B. V., Holmes, D. L., Guth, M., & Kovac, P. (1979). The potential of children as eyewitnesses. *Law and Human Behavior, 3,* 295-305.

Maryland v. *Craig,* 110 S. Ct. 3157 (1990).

Mason, M. A. (1991, Winter-Fall). A judicial dilemma: expert witness testimony in child sex abuse cases. *Psychiatry and Law,* 185-219.

McGough, L. (in press). *Fragile voices: The child witness in American courts.* New Haven, CT: Yale University Press.

Melton, G. (1992). Children as partners for justice: Next steps for developmentalists. *Monographs of the Society for Research in Child Development,* 57(Serial No. 229), 153-159.

Messerschmidt, R. (1933). The suggestibility of boys and girls between the ages of six and sixteen. *Journal of Genetic Psychology, 43*, 422-437.

Montoya, J. (1992). On truth and shielding in child abuse trials. *Hastings Law Journal, 43*, 1259-1319.

Montoya, J. (1993). *When fragile voices intersect with a fragile process: Pretrial interrogation of child witnesses.* Unpublished monograph. University of San Diego School of Law.

Moss, D. C. (1988). "Real" dolls too suggestive: Do anatomically correct dolls lead to false abuse charges? *American Bar Association Journal, 24*, 24-25.

Myers, J. E. (1992). *Legal issues in child abuse and neglect.* Newbury Park, CA: Sage.

Myers, J. E. (1993). Expert testimony regarding child sexual abuse. *Child Abuse and Neglect, 17*, 175-185.

Nathan, D. (1987). The making of a modern witch trial. *The Village Voice, 33, 19-32.*

National Center for Child Abuse and Neglect. (1988). *Study of national incidence and prevalence of child abuse and neglect: 1988.* Gaithersburg, MD: U.S. Department of Health and Human Services.

National Center for Child Abuse and Neglect. (1993). *National Child Abuse and Neglect Data System, 1991: Summary data component.* Gaithersburg, MD: U.S. Department of Health and Human Services.

New Jersey v. Michaels (1993). 264 N.J. super (AppDiv) 579.

Oates, K., & Shrimpton, S. (1991). Children's memories for stressful and nonstressful events. *Medicine, Science, and the Law, 31*, 4-10.

Ornstein, P. A., Baker-Ward, L., Gordon, B., & Merritt, K. (1993). *Children's memory for medical procedures.* Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans.

Ornstein, P. A., Gordon, B. N., & Larus, D. (1992). Children's memory for a personally experienced event: Implications for testimony. *Applied Cognitive Psychology, 6*, 49-60.

Otis, M. (1924). A study of suggestibility in children. *Archives of Psychology, 11*, 5-108.

Peters, D. D., Wyatt, G. E., & Finkelhor, D. (1986). Prevalence. In D. Finkelhor (Ed.), *A sourcebook on child sexual abuse.* Beverly Hills: Sage.

Popper, K. R. (1962). *Conjectures and reflections.* New York: Basic Books.

Rabinowitz, D. (1990, May). From the mouths of babes to a jail cell: Child abuse and the abuse of justice. *Harper's Magazine*, pp. 52-63.

Raskin, D., & Esplin, P. (1991). Assessment of children's statements of sexual abuse. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 153-164). Washington, DC: American Psychological Association.

Raskin, D. & Yuille, J. (1989). Problems in evaluating interviews of children in sexual abuse cases. In S. J. Ceci, M. P. Toglia, & D. F. Ross (Eds.), *Adults' perceptions of children's testimony* (pp. 184-207). New York: Springer-Verlag.

Robin, M. (1991). The social construction of child abuse and "false allegations." *Child and Youth Services, 15*, 1-34.

Rudy, L., & Goodman, G. S. (1991). Effects of participation on children's reports: Implications for children's testimony. *Developmental Psychology, 27*, 527-538.

Sauer, M. (1993, August 29). Decade of accusations. *San Diego Union Tribune*, D1-D3.

Saywitz, K. J., Goodman, G. S., Nicholas, E., & Moan, S. F. (1991). Children's memories of a physical examination involving genital touch: Implications for reports of child sexual abuse. *Journal of Consulting and Clinical Psychology, 59*, 682-691.

Schooler, J. W., Gerhard, D., & Loftus, E. F. (1986). Qualities of the unreal. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 171-181.

Sgroi, S. M. (1982). *Handbook of clinical intervention in child sexual abuse.* Lexington, MA: Lexington Books.

Shaw, D. (1990, January 19-22, Series). McMartin verdict: Not guilty. *Los Angeles Times.*

Sherman, I. (1925). The suggestibility of normal and mentally defective children. *Comparative Psychology Monographs, 2.*

Siegel, J. M., Sorenson, S. B., Golding, J. M., Burnam, M. A., & Stein, J. A. (1987). The prevalence of childhood sexual assault: The Los Angeles epidemiologic catchment area project. *American Journal of Epidemiology, 126*, 1141-1153.

Spencer, J. R. (1992). Court experts and expert witnesses. In R. W. Rideout & B. Hepple (Eds.), *Current legal problems, Vol. 45* (pp. 214-236). London: Oxford University Press.

Steller, M. (1991). Commentary: Rehabilitation of the child witness. In J. L. Doris (Ed.), *The suggestibility of children's recollections* (pp. 106-109). Washington, DC: American Psychological Association.

Stern, W. (1910). Abstracts of lectures on the psychology of testimony and on the study of individuality. *American Journal of Psychology, 21*, 270-282.

Summit, R. (1983). The child sexual abuse accommodation syndrome. *Child Abuse and Neglect, 7*, 177-193.

Taylor, J. P. (cited in S. R. Gross) (1858). *Treatise on the law of evidence, 3rd Ed.* §§ 45-50, 65-69.

Terr, L. (1988). Anatomically correct dolls: Should they be used as a basis for expert testimony? *Journal of the American Academy of Child and Adolescent Psychiatry, 27*, 254-257.

Varendonck, J. (1911). Les témoignages d'enfants dans un proces retentissant [The testimony of children in a famous trial]. *Archives de psychologie, 11*, 129-171.

Wolfner, G., Faust, D., & Dawes, R. (1993). The use of anatomical dolls in sexual abuse evaluations: The state of the science. *Applied and Preventative Psychology, 2*, 1-11.

Wyatt, G.E. (1985). The sexual abuse of Afro-American and white American women in childhood. *Child Abuse and Neglect, 9*, 507-519.

About the Authors

Steve Ceci, Ph.D., is Helen L. Carr Professor of Developmental Psychology in the Department of Human Development and Family Studies, Cornell University.

Maggie Bruck, Ph.D., is an Associate Professor in the departments of Psychology and Pediatrics at McGill University.

Acknowledgements

This research was supported, in part, by grants to the first author from the National Institute of Child Health and Human Development and to the second author by a grant from the Natural Sciences and Engineering Research Council.

Next issue: (Winter 1993). Integrating science and ethics in research with high-risk and underserved children and youth [tentative title]. Celia B. Fisher.

Subscriptions ($10 to nonmembers of SRCD) and single issues ($3.50) are available from:

SRCD Executive Office
University of Michigan
300 North Ingalls, 10th floor
Ann Arbor, MI 48109-0406

Social Policy Report *is a quarterly publication of the Committee on Child Development, Public Policy and Public Information of the Society for Research in Child Development. Copyright of the articles published in the* Report *is maintained by the originator[s] of the work. Statements appearing in the* Report *are the views of the author and do not imply endorsement by the Editor or by SRCD.*

Editor: Nancy G. Thomas
Phone: (303) 925-5516
Fax: (303) 925-9570

72

SRCD Executive Office
University of Michigan
300 N. Ingalls 10th Floor
Ann Arbor, MI 48109-0406

Non Profit
Organization
U.S. Postage
PAID
Permit No. 587
Ann Arbor, Michigan

SOCIAL POLICY REPORT

SOCIAL POLICY REPORT
Society For Research In Child Development

Volume VII, Number 4 Winter 1993

INTEGRATING SCIENCE AND ETHICS IN RESEARCH WITH HIGH-RISK CHILDREN AND YOUTH

Celia B. Fisher

Background

Over the last two decades child development research has reflected an ever growing concern with the daily challenges facing vulnerable and socially disenfranchised children and their families (Fisher & Lerner, 1994; Fisher et al., 1993; Fisher & Tryon, 1988, 1990). As developmentalists seek to apply the scientific method to describe, explain, and enhance the status of disadvantaged children and youth, they are encountering ethical dilemmas to which current guidelines offer incomplete answers (Fisher & Rosendahl, 1990). The scientific and ethical dimensions of this research often appear to have mutually exclusive goals. Whereas the science dimension involves a search for truth through experimental controls, the ethics dimension is directed toward the protection of participant autonomy and welfare through means that often appear to jeopardize controls. Researchers must confront questions such as

- how to balance the need to secure a valid sample with the need to protect the autonomy of potential subjects who may need assistance or be especially vulnerable;
- how to handle researcher-participant confidentiality when problems such as illicit or abusive circumstances are uncovered in the course of conducting research;
- how to take an active part in the dissemination and application of research findings in ways that avoid a misinterpretation or overgeneralization of data that may jeopardize the development of sound public policy and programs.

PS 021022

Traditionally trained developmentalists are findings themselves unprepared to resolve dilemmas like these. As a consequence, the integrity and adequacy of ethical directives are at the forefront of deliberations about scientific practices as they apply to high-risk children and youth.

Federal and Professional Guidelines for Research with Children and Youth

Formal guidelines for ethical practice in the conduct of research with human subjects have their origin in the *Nuremberg Code* (1946), a set of international guidelines developed in response to the atrocities of Nazi medical experimentation. Since then, federal agencies and professional organizations have developed ethical codes aimed at providing standards of conduct for research with adults and minors (e.g., *Protection of Human Subjects*, Department of Health and Human Services [DHHS], 1991; *Ethical Principles of Psychologists and Code of Conduct*, American Psychological Association, 1992; *Ethical Standards of the American Educational Research Association*, 1992; and *Ethical Standards for Research with Children*, Society for Research in Child Development [SRCD], 1993).

One of the most detailed ethical guidelines for behavioral research, often referred to as the *Belmont Report*, was presented by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (Department of Health, Education, and Welfare [DHEW], 1978). In addition to suggesting specific ethical standards for human subjects research, the report articulated three moral principles upon which ethical conduct should be based: *beneficence, respect*, and *justice*. These principles underscore the researcher's obligation to protect participant welfare, to respect participant autonomy and privacy, and to ensure the equitable distribution of benefits and costs associated with research participation. Today these three moral values are considered the foundation of ethical decision making in research with children and adolescents (e.g., Beauchamp & Childress, 1979; Fisher & Tryon, 1990; Gaylin & Macklin, 1982; Koocher & Keith-Spiegel, 1990).

Beneficence. The principle of *beneficence* is embodied in the maximization of possible benefits and minimization of possible harm (DHEW, 1978). Beneficial outcomes can be accomplished through experimentation directly or indirectly. For example, research designed to evaluate the efficacy of a developmental intervention has the potential to benefit *directly* the child or adolescent assigned to the "treatment" group. On the other hand, a study designed to describe the developmental consequences of high-risk conditions of childhood and adolescence has the potential to benefit *indirectly* the participant's social group through future application of research findings by parents, practitioners, or policymakers.

Research can also expose participants directly or indirectly to experiences that result in serious or permanent harm. For example, parents of children assigned to a "no treatment" control group may forego service options for their children available outside the experimental setting (Fisher, 1991; Fisher & Rosendahl, 1990). Alternatively, experimental procedures or assessment strategies may inadvertently elicit stress in child and adolescent research participants. It is because of such contingencies that most professional codes and federal regulations require investigators to weigh the potential balance of harms and benefits to participants before deciding to conduct research (e.g., see Principle 1, SRCD, 1993). At present, however, researchers, institutional review boards (IRBs), and policymakers have little guidance in determining the acceptable risk and benefit balances.

Respect. The principle of *respect* reflects a moral concern that individuals be treated as autonomous agents and that more vulnerable persons with diminished autonomy have these rights protected (DHEW, 1978). This principle is reflected in professional codes and federal regulations that require that potential subjects be fully informed about the nature of the research, what will be required of them, the potential risks and benefits of participation, and their right to withhold or withdraw consent without penalty (Faden & Beauchamp, 1986; see Principle 2, SRCD, 1993).

Consent procedures for research participants who are minors require special consideration for several reasons. Children and adolescents do not have the legal capacity to consent (Melton, Koocher, & Saks, 1983). They may lack the cognitive capacity, depending on the age of the child and complexity of the situation, to comprehend the rationale or nature of the experimental procedures (see Thompson, 1990). And they may perceive they lack power to refuse participation (Fisher & Rosendahl, 1990; Keith-Spiegel, 1983; Koocher & Keith-Spiegel, 1990; Levine, 1986). In light of these issues, federal regulations and professional codes (e.g., see Principle 3, SRCD, 1993) require the informed consent of parents, legal guardians, or those who act *in loco parentis* before a child can participate in research. In addition, special child assent procedures are recommended for children older than 7. According to the SRCD *Ethical Standards for Research with Children*, "Assent means that the child shows some form of agreement to participate without necessarily comprehending the full significance of the research necessary to give informed consent" (see Principle 2, SRCD, 1993, p. 337). The objection of a child of any age is binding unless the experimental treatment holds out the prospect of direct benefit that is important to the child and only achievable through the research.

Justice. The principle of *justice* refers to the moral value placed on equitable distribution of social benefits and costs (Rawls, 1971). This principle, as articulated in federal and professional codes for human subjects research, requires equitable selection and treatment of participants regardless of nationality, race, age, sex, or social status. Justice is expressed in the research design of developmental studies when investigators manage to maintain a balance between including diverse populations and remaining sensitive to the possibility of differential impacts. From this perspective, it is important to include disadvantaged and minority populations in developmental research, since not to do so runs the risk of depriving society of a knowledge base from which suitable and effective services can be constructed (Fisher, 1991; Fisher & Brennan, 1992; Graham, 1992; Hagen, Paul, Gibb, & Wolters, 1990; McLoyd & Randolph, 1984).

Overview

The principles of beneficence, respect, and justice can serve as valuable guides for researchers who work with vulnerable populations. This report explores the implications of these principles for such research in all its phases—recruitment, implementation, and the dissemination of findings. The first section provides an overview of the ethical challenges associated with informed consent procedures for research with children and youth. Included are an exploration of the informed, rational, and voluntary requirements of consent to participate in descriptive and intervention research, and of the conflicts that arise when guardian consent is unavailable or not in the best interests of the child. The second section addresses the investigator's responsibilities when problems that may jeopardize the minor participant's well-being are discovered during the course of research. The researcher must consider the validity of risk estimates, how and with whom information should be shared, and the federal and state laws regulating reporting responsibilities. Finally, the third section deals with the ethical concerns associated with generalizing the results of research based on large social units to individuals and diverse populations. Included is discussion of the ethical implications of overestimating the applied relevance of developmental research, of failing to consider individual and cultural differences when designing and evaluating social programs, and of applying the developmental knowledge base to legal decisions relevant to the welfare of children and youth. The report concludes with recommendations for future policies.

Securing the Sample: Informed Consent for Participation in Descriptive and Intervention Research

Informed consent, as currently required by federal and professional ethical guidelines, is seen by many as the primary means of protecting participants. Consent procedures are designed to protect participant autonomy and self-determination by ensuring that the decision to participate is informed, rational, and voluntary (Freedman, 1975). Participant rights are further safeguarded in the case of children and adolescents through the requirement of parental or guardian consent. In addition, to ensure that the rights and privacy of all persons involved in the research are protected, informed consent is obtained from all involved individuals. A teacher's consent may be required, for example, when the focus of an investigation is the student-teacher interaction (see Principle 4, SRCD, 1993).

The Informed Requirement of Consent

The *informed* aspect of consent refers to the requirement that prospective participants be provided with information regarding all procedures that might influence their willingness to participate themselves or allow their child to participate. Such information includes a description of the procedures, the time required to participate, the profession and institutional affiliation of the investigator(s), foreseeable risks or benefits of participation, the extent as well as limits of confidentiality, the voluntary nature of participation, and procedures for obtaining results and conclusions of the research. For both guardian consent and child assent, the informing procedure should be appropriate to the language usage and level of comprehension of parent and child (Fisher & Rosendahl, 1990; Thompson, 1990). In addition, the researcher should answer all questions posed by participants or their guardians.

The degree to which parental consent is truly informed can be affected by the specific research question. For example, in a study exploring the developmental correlates of maternal depression, the investigator may need to take into account the potential impact of symptoms associated with depression, such as difficulties in concentration or feelings of worthlessness, on the mother's capacity to understand the nature of the study and to respond in her own best interest and that of her child.

The Rational Nature of Consent

The *rational* requirement of consent procedures reflects an acknowledgement that the context in which parents of children at risk are recruited may work against a rational approach to consent. The very conditions that identify the child as a potential participant may make it difficult for parents to recognize the limited or nonexistent benefits of participation in research meant to be descriptive rather than intervening.

For example, a family with a child who exhibits a pattern of disabilities but who has no definitive diagnosis or treatment may look to any contact with professionals as potential assistance; and the family may assume erroneously that some benefit will come of participation (Carroll, Schneider, & Wesley, 1985; Fisher & Rosendahl, 1990). Members of vulnerable families recruited to assess the efficacy of a developmental intervention may mistake descriptions of potential benefits of participation for presumed benefits. Thus, recruitment procedures for participation in such research must ensure that participants and/or guardians fully understand the nature and risks of participation, including the potential consequences of random assignment to a control condition that has been empirically associated with later developmental problems.

Investigators studying development in low-risk populations must also be sensitive to unintentional violations of the rational aspect of informed consent. For example, a procedure may itself have potentially stressful consequences (Fisher & Rosendahl, 1990). As an illustration, consider a study designed to follow the development of children with social problems that includes a screening of kindergartners to identify those high in aggressive behaviors or rejected by peers. Let us say then that these "aggressive" and "rejected" kindergartners are to be the focus of a second phase of the study. Some parents of the "aggressive" children selected for this second phase may become concerned that their child has "special" problems, and they may misperceive participation in the research as a means of helping their child or of providing ongoing "diagnosis" of the child's "condition." Informed consent procedures that draw upon our knowledge of affective factors associated with parenting can help to ensure that guardian consent is both informed and rational.

The Voluntary Aspect of Informed Consent

Informing participants. Investigators who study the development of disenfranchised children and their families must be particularly sensitive to potential violations of the voluntary aspect of guardian consent (Fisher & Rosendahl, 1990). Vulnerable families, contacted while seeking services at community mental health or medical facilities, may be concerned that failure to consent will result in a discontinuation of services for themselves or for their children (Fisher, 1991). Under these circumstances, special care needs to be taken to clarify that the right to services is not contingent on consent, and that the child can withdraw even after consent has been given.

Some participants and their guardians recruited for intervention research may prefer either the treatment or matched control condition to random assignment. The traditional elimination of these individuals from the subject pool may compromise the voluntary nature of informed consent. Some individuals desperate to help themselves or their children may consent to random assignment because they see the chance of receiving services as preferable to receiving no services at all. This procedure may also limit the generalizability of findings to families who are willing to be randomly assigned (Fisher, 1991; Fisher & Rosendahl, 1990). For example, adolescent mothers who prefer nonrandomized assignment to the treatment group of a parenting skills intervention study may have the personal resources to benefit most from this type of intervention. A study finding no treatment effects for young women willing to be randomized might lead to the erroneous conclusion that such a program was ineffective for all teenage mothers. A novel approach to this problem, the *semi-randomized clinical trial* (Veatch, 1987), assigns participants to four groups: two groups comprised of individuals who have agreed to random assignment to either the treatment or control condition, and two groups who have selected nonrandom assignment to the two conditions. This design respects participant autonomy and enables researchers to compare the responses of participants who agree to randomization with those who do not.

Incentives. To ensure that parental judgments are rational, voluntary, and uncoerced, investigators studying at-risk populations need to take special precautions when offering inducements for research participation, on the grounds that undue incentives can invalidate informed consent. The decision to provide inducements creates a tension between compensating individuals fairly for their time and inconvenience and subjecting them to undue coercion to participate in procedures to which they might not otherwise consent. This is particularly true in the case of the recruitment of impoverished persons, who, because they need the income, may be willing to assume, or have their children assume, extraordinary burdens (Levine, 1986). Unfortunately, at present there is little consensus on what defines due and undue incentives for research participation (Macklin, 1981).

In making decisions about the use of incentives, the investigator must consider the ethical principle of justice, that is, how the particular inducement will affect the equitable distribution of both the burdens and benefits of research participation. Ideally, offering incentives should enable a balanced sampling of individuals from all socioeconomic levels. However, the varying economic circumstances of potential participants will lead to varying perceptions of a cash inducement. One response to this dilemma has been to take a market approach to incentives, whereby participation is viewed as an ordinary job requiring relatively unskilled labor, and payment is determined at a level sufficient to attract the desirable number and diversity of research participants (Levine, 1986).

Offering incentives for child and adolescent participation requires additional consideration. For example, establishing procedures whereby cash payments or services go directly to the minor participant rather than the guardian has been one way to ensure that participants and their guardians see the child's efforts as fairly compensated. An additional safeguard is to offer inducements that do "not unduly exceed the range of incentives that the child normally experiences" (see Principle 5, SRCD, 1993, p. 337).

The Best Interests of the Child

Requiring guardian consent assumes that the child comes from a reasonably secure family setting in which the child and his or her parents share loving relationships (Gaylin & Macklin, 1982; Levine, 1986). However, significant numbers of children from high-risk physical or social environments have undetermined custody, nonrelative guardians, or state guardianship (Gibbs, 1990; Hendren, 1991). In addition, high-risk physical and social conditions may in themselves make obtaining consent from identified guardians difficult. In the absence of specific consent policies regarding the "best interests of the child," developmental investigators are torn between excluding children with uncertain custody or waiving parental consent and the participant protection such consent is designed to afford.

Passive consent. Difficulties in acquiring guardian consent from at-risk populations has led in recent years to renewed debate over the use of passive guardian consent. Passive consent is the procedure by which guardians are sent forms describing the research and are asked to respond *only* if they do *not* wish their child to participate. This procedure does not meet the criterion of ethical practice implied by the principle of respect for self-governance, in that the researcher cannot know for certain that the child's guardian received the information or that the guardian's failure to respond reflects agreement to have the child participate. Recently, investigators have drawn upon the principle of beneficence to argue that since active consent procedures produce lower rates of return for low-income, disadvantaged youth, passive consent procedures are justified to ensure that these populations will receive the benefits of knowledge derived from scientific inquiry. The principle of justice points, however, to the ethical vulnerability of this argument; the use of passive consent with disadvantaged populations means that children of low-economic status, in contrast to children from middle-income homes, are not receiving the equal protection afforded by guardian consent.

Underlying the advocacy for passive consent is an implicit assumption that a caring and knowledgeable guardian would perceive the research as important and as desirable for his or her child. Equally unsupported, but more unfortunate, is the additional inference that disadvantaged parents who do not return consent forms either lack the knowledge to appreciate the importance of the research or are unconcerned about their child's welfare. No empirical data exist to support either of these assumptions, nor is there research comparing the response rates for middle-class and low-income parents for studies that would have comparable meaning to the lives of their children. In fact, many people from low-income, ethnic minority communities may look upon research projects with suspicion. The IQ debate and associated tracking movements in schools, the scandal surrounding the Tuskegee syphilis study, and the brewing scandal over government radiation experiments in the 1940s through the 1960s serve to undermine trust in social scientists as guardians of ethical treatment when prospective subjects are minorities (J. F. Jackson, personal communication, January 6, 1994). Given the paucity of data and the potential threats of passive consent to participant autonomy and privacy, these unwarranted assumptions risk substituting investigator paternalism for parental permission. Consequently, an empirical assessment of who fails to consent when active consent is used with different populations of high-risk youth, and how we can develop more effective recruitment procedures to increase active guardian consent, should be an important part of the nation's child and adolescent research agenda.

Can adolescents participate in research without guardian consent? The social challenges facing today's youth raise questions concerning the rights of minors to consent to research participation without the knowledge or permission of their legal guardians. For example, the requirement that guardians give consent for the research participation of adolescents in studies on substance abuse or sexual activity, may, rather than protect participant rights, violate a teenager's privacy or even jeopardize his or her welfare (Brooks-Gunn & Rotheram-Borus, in press).

In recent years the legal-medical model defining the rights of adolescents to consent to medical interventions has provided a framework for evaluating the conditions under which guardian consent for research participation might be waived (Holder, 1981). All states currently allow minors to consent independently to treatment for venereal disease, and states specify statutory ages (usually 15 or 16) at which teenagers may consent to medical procedures (including, in some states, alcohol or drug abuse treatment) without parental involvement (Holder, 1981). In that the law grants teenagers the autonomy to make decisions concerning their medical treatment in such instances, it seems both reasonable and ethical to grant them the same autonomy to decide whether or not to participate in research that examines their reasons for or reactions to seeking these treatments (Scarr, in press). Moreover, each state has statutes defining the conditions of emancipation, such as adolescent parenthood or independent living arrangements, which can be used as guides for ethical decisions about the rights of minors to give autonomous consent for research participation. Developmentalists should nevertheless guard against simply substituting such statutes for more principled decisions about guardian consent waivers based on developmental considerations.

Federal guidelines for research with children (DHHS, 1983) allow an institutional review board (IRB) to determine the conditions under which requiring parental consent is not a reasonable protection of minor participants. In so determining, the IRB can waive parental permission if appropriate mechanisms for protecting the child are substituted. It would seem, based on this principle, that guardian consent would be waived—for instance, in the case of teenage runaways—if procedures are established to ensure that the adolescent's own consent is informed, rational, and voluntary. Researchers must be mindful, however, of Robert Veatch's (1981) cautionary comments that guardian consent should not be waived simply to facilitate the research process. Decisions regarding whether adolescents should participate in research without parental consent should be based on the potential benefits to the participant rather than the utility needs of the researcher.

Problems Identified During the Course of Research: Sharing, Referring, and Reporting

In the course of conducting their studies with high-risk children, researchers may utilize assessment instruments or specific methodologies that yield information indicating that a research participant has previously undetected cognitive, socioemotional, or physical problems. Under these circumstances the principle of beneficence would lead the investigator to consider whether to share such information with participants and/or their parents, to provide information about referral services, or, under extreme circumstances, to report the problem to local authorities (Fisher, in press; Fisher & Rosendahl, 1990).

Developmental researchers have traditionally been reluctant to communicate the results of individual assessments to parents, both out of concern for confidentiality and a healthy skepticism regarding the clinical relevance of research-derived information (Asher, 1993; Brooks-Gunn & Rotheram-Borus, in press; Fisher & Brennan, 1992). In fact, one recent survey of leading developmental journals found that less than 1% of the empirical articles mentioned that information was shared with parents or that referrals were made for developmental problems discovered during the course of research (Fisher & Brennan, 1992).

According to Principle 9 of the SRCD *Ethical Standards for Research with Children,*

> When, in the course of research, information comes to the investigator's attention that may jeopardize the child's well-being, the investigator has a responsibility to discuss the information with the parents or guardians and with those expert in the field in order that they may arrange the necessary assistance. (SRCD, 1993, p. 338)

In deciding whether to share research-derived information about child and adolescent participants, the investigator must consider the validity of the risk estimates, how and with whom information should be shared, and the federal and state laws which regulate the reporting responsibilities of researchers working with minors.

The Validity of Risk Estimates

Any decision to share information obtained during the course of research must include an initial consideration of whether the experimental procedures provide a valid means of identifying the presence of specific developmental disorders or risks (Fisher, 1991; in press). Investigators must guard against underestimating or overestimating the predictive utility of developmental measures and should proceed with caution when making evaluative statements or giving advice (see Principle 13, SRCD, 1993).

Information derived from diagnostic measures. Investigators, in examining the development of children of vulnerable populations, often utilize standardized measures of individual psychological functioning. The Wechsler Intelligence Scale for Children ([WISC-R], Psychological Corporation, 1991), the Child Behavior Checklist ([CBCL], Achenbach & Edelbrock, 1983), and the Children's Depression Inventory ([CDI], Kovacs, 1981) are commonly used to assess the cognitive or emotional status of child and adolescent research participants. To the extent that these assessments have diagnostic validity, the researcher may come to possess information about potential conditions that jeopardize the child's well-being. For example, a pattern of WISC-R responses may suggest a learning disability; the CDI may yield information indicating childhood depression; and maternal reports on the CBCL may indicate anxiety disorders of childhood or adolescence. When this happens, investigators may bear a responsibility to share information in order that the child obtain needed assistance (see Principle 9, SRCD, 1993).

Information derived from descriptive measures. Some assessment instruments that are valid for identifying group differences in the experimental setting may not have the psychometric properties necessary for a meaningful assessment of individual abilities or disabilities (Fisher, in press; Fisher & Brennan, 1992). Sharing with parents such research-derived observations about a minor participant runs the risk of creating stressful or harmful misconceptions about the child's development.

For example, while adolescent experimentation with alcohol has been associated with depression, school problems, delinquency, and unprotected sex (Kandel, Davies, Karus, & Yamaguchi, 1986), we do not yet have sufficient information about either the degree of alcohol consumption at which such effects are seen or the specific personality characteristics or social factors involved to determine whether a particular adolescent will progress to abuse of alcohol during the teenage years or later in life (Pentz, 1994). Given the lack of empirical evidence documenting individual risk for teenage alcohol use, sharing information on adolescent drinking behaviors with parents or discussing the implications of specific responses with the adolescents themselves could be potentially misleading and outside the boundaries of good ethical practice for developmental scientists. On the other hand, as outlined in the next section, in the absence of diagnostically valid measures of individual functioning, researchers may wish to incorporate more general information about psychological or social services into their standard debriefing procedures (Fisher & Rosendahl, 1990).

Clarifying the Limits of Communication and Referral Procedures

Once an investigator has determined the direct or indirect relevance of empirical observations to participant welfare, she or he must decide what type of information will be shared and with whom. Advance planning is the best means of establishing communication and referral procedures that will be well received by participants and their guardians and can be successfully implemented.

Clarifying the extent and limits of communications as part of consent procedures can help establish realistic participant and guardian expectations about information the investigator can and cannot provide (Fisher, in press). For example, researchers should not promise children or adolescents confidentiality if they are prepared to share some portion of the child's experimental responses with parents (Asher, 1993; Brooks-Gunn & Rotheram-Borus, in press). In contrast, if information is not to be shared, investigators should not give parents the impression that participation in the research will advance their knowledge about their child. As noted in the previous section, specifying the limits of information that can be shared helps ensure that guardian consent is both informed and rational.

Providing Assessment Results

When an investigator has determined that it would be to a child's advantage to share valid diagnostic information with parents, she or he must ensure that test administration, interpretation, and communication of results meet current standards of ethical practice. Research personnel must be competent to administer and score tests appropriately. In interpreting results they must take care to identify situations in which particular assessment norms may not be applicable to individual diagnoses (e.g., for children from language communities or cultural groups not included in the population on which the test was standardized).

Research personnel must also be trained to explain the assessment results to parents in language that is understandable. Moreover, if information is to be shared with individuals other than the child's guardian (e.g., the child's pediatrician or academic advisor), the investigator should ensure that these individuals are qualified to use and interpret psychological assessments. Information should never be shared with other professionals without the explicit, written consent of the guardian (Fisher, in press).

Providing Referral Information

In recent years, developmental researchers have generated information on the risk and protective factors associated with a wide range of child and adolescent behaviors. For example, we know that peer rejection is associated with loneliness in elementary school children (Asher & Dodge, 1986), that low self-esteem is correlated with maladjustment, depression, and anxiety in both school and peer relations (Damon, 1983), and that adolescent pregnancy is a marker variable for a cluster of developmental problems and risk factors facing young women (Chase-Lansdale & Brooks-Gunn, 1994).

Although the lack of established diagnostic assessments based on this knowledge base preclude sharing the results of individual performance, investigators working in these areas can provide participants or their guardians information about existing services relevant to the developmental issues under investigation (Fisher & Rosendahl, 1990). For example, an investigator studying the development of children's self-esteem can inform parents about the range of behaviors associated with positive and negative self-concepts, and provide a list of community agencies offering psychosocial evaluations and treatment of developmental problems. Care should be taken, however, that parents understand that the investigator is not recommending assessment or treatment, but merely providing information about such services, should the parent be concerned about his or her child's development.

Identifying such services, especially low-cost services, may require the investigator to do some "detective" work within the professional community. The difficulties encountered in finding appropriate service systems for children and adolescents can help the investigator understand and appreciate participants' service options, thereby enriching future research designs focused on the interface between developmental risk and community supports.

Working with practitioners. In some situations the researcher may be in a position to identify specific professionals to assist participants with serious problems. For example, a developmentalist studying the correlates of adolescent pregnancy may establish a relationship with a medical facility that can provide prenatal or postnatal care to teenage research participants. Referrals to such services can be made for participants who lack medical care. An investigator can also collaborate with mental health professionals to develop referral procedures for participants in need of psychological services, including crisis intervention in the case of the participant whose responses to the experimental procedure indicate a grave situation or life-threatening behaviors calling for immediate attention (Fisher, 1993). Investigators can enlist practitioners during the study's design phase to help develop guidelines for identifying participant problems and for handling consultation with participants as situations arise (Fisher & Rosendahl, 1990).

Allowing for self-referral. Asher (1993) has developed and tested an innovative procedure of *self-referral* devised for peer-rejected and lonely children recruited as part of normative samples in the schools. With the cooperation of school administrators and staff, Asher gave all child participants the opportunity to check a box if they wanted to speak to the school social worker about "things that bother you," "questions you might have about things you want to change," or "learning how to make friends." Asher's finding that low-accepted children were more likely to self-refer than other children strongly supports the proposal that self-referral is an ethical procedure that can help high-risk children seek assistance. Continued refinement of this procedure will help identify conditions under which self-referrals should or should not require parental consent (Scarr, in press).

Providing treatment for control group participants. Intervention research poses additional ethical problems in that the standard treatment-control group design intentionally denies services to children and youth with identified developmental risk. There is a growing and persuasive literature indicating that basic services, such as nutrition supplements, medical care, and education, can help overcome biological and environmental risks (e.g., Chase-Lansdale & Brooks-Gunn, 1994; Wasik, Ramey, Bryant, & Sparling, 1990). This raises ethical concerns about the use of *no-treatment* control groups for the evaluation of new intervention strategies in the face of empirical evidence that such basic services will benefit participants (Fisher & Rosendahl, 1990). Providing control groups with basic services can jeopardize conclusions that the treatment *per se* is significantly better than no treatment at all. But comparing psychosocial intervention to less costly basic health and social services may be the more honest intellectual question, especially in times of limited economic resources (Fisher, 1991).

When no prior treatment strategies have been established, a primary ethical consideration is whether to provide services for control group participants after the experimental intervention has been evaluated. At the completion of clinical trials research aimed at assessing the efficacy of short-term intervention strategies, it is standard ethical practice to offer adult control groups the treatment found to be most effective. However, at the conclusion of long-term studies designed to evaluate the effect of an early intervention on later development, the treatment may be developmentally inappropriate for the (now older) controls. One way to address this issue is to utilize postintervention assessments for referrals to appropriate community services (Fisher, 1991).

Reporting Illegal and Abusive Behaviors

Investigators working with high-risk populations may, during the course of research, become privy to information suggesting illegal conduct, use of addictive products, or threats of harm or violence to the participant or identified others. Researchers need to be aware that certain decisions—for example, whether to report to authorities suspected physical or sexual abuse of child participants, or to disclose to the court research-derived information about illegal participant activities—are not only ethical decisions, but in many states have legal implications as well (Appelbaum & Rosenbaum, 1989; Fisher, 1991; Hoagwood, in press; Liss, in press).

Child abuse and neglect. In 1976, federal law (the Child Abuse Prevention and Treatment Act) began to require that states enact statutes to mandate the reporting of suspected child abuse or neglect. In an extensive review of the 50 different state statutes on child abuse and neglect reporting, Liss (in press) documented that in at least 13 states, researchers, as members of the general citizenry, are mandated reporters. Other states vary as to whether researchers are included in the list of mandated reporters, and, if included, whether reporting is required if problematic information is obtained outside the research protocol. As Liss points out, until case law decisions can provide more guidance, researchers must rely on reviewing their own state laws to determine their personal responsibilities in child abuse and neglect reporting. Developmental scientists engaged in multidisciplinary collaborations should also determine whether their colleagues (e.g., pediatricians, school psychologists) are mandated to report abuse, because whether they are or not has implications for researcher-participant confidentiality and the protection of information discovered during the course of research (Fisher, 1991). Having determined their own reporting responsibilities, investigators must also ensure that research assistants are trained to recognize indicators of abuse and to follow appropriate procedures. Finally, the limits of confidentiality, as determined by state reporting laws, should be fully communicated to both parents and minor participants at the time of informed consent.

Violent behavior. A recent concern of mental health researchers is whether they share with their practitioner colleagues the responsibility to prevent the occurrence of violent behavior by a research subject (Appelbaum & Rosenbaum, 1989). This issue is particularly relevant to developmental scientists who study family violence or delinquency.

In 1969, Prosenjit Poddar, a student at the University of California Berkeley who had been a client at the university student health center, murdered Tatiana Tarasoff. Prior to the murder, the psychologist counseling Poddar had notified the campus police that his client was dangerous, but he withdrew a letter recommending confinement after the police concluded Poddar posed no harm. Following a series of appeals by the Tarasoff family, the Supreme Court of California in *Tarasoff v. Regents of the University of California* (1976) concluded that the therapist had a duty to warn third parties about dangers posed by clients. In response to this decision, several states have held that a psychotherapist has a duty to protect potential victims from a patient's violent acts if the therapist (1) has a "special relationship" with the patient, (2) can predict that violence will occur, and (3) can identify the potential victim. While no court case has yet applied this duty-to-protect to the research context, given appropriate circumstances (e.g., the special relationship, prediction of violence, and victim identification), some researchers working with violence-prone populations may in the future be expected to assume the same duty (Appelbaum & Rosenbaum, 1989).

The Certificate of Confidentiality. Developmental scientists may also find themselves in the position of collecting information about *past* abuse, violence, or illegal activities. The researcher needs to be aware, and must inform his or her research participants, that confidential research records are subject to subpoena (Melton, 1990a, 1990b; see Principle 11, SRCD, 1993). Clearly, such legal vulnerability creates barriers to participant responding and recruiting; potential participants are unlikely to agree to participate or answer honestly if they believe their responses will result in legal proceedings against them. To protect the privacy and identity of research participants, investigators can apply for a Certificate of Confidentiality issued by the Department of Health and Human Services (Hoagwood, in press). This certificate protects the researcher against legally compelled disclosure of identifying information about individual research participants. However, the certificate does not override all reporting requirements. While it can protect the privacy of participant reports of sexual activities, substance abuse, illegal conduct, psychological well-being, and past violent or abuse acts, it may not override state and local laws governing the reporting of current child abuse or communicable disease. From an ethical standpoint, researchers who have been granted the certificate should inform their participants about the extent and limitations of its protection and plan further ways to protect the confidentiality of their data (Hoagwood, in press).

Caveats for Sharing, Referring, and Reporting

This section has addressed the issues researchers must consider when they obtain information suggesting that participants need assistance. However, as Scarr (in press) has pointed out, researchers must be careful not to go beyond their areas of competence or to sacrifice their primary responsibility to the research by offering referrals, when little is known about how to remediate or prevent the developmental problems under investigation.

Developmental scientists must also be wary about overreporting or overreferring children and adolescents from ethnic minority and low-income families (Edelbrock, 1994; Scott-Jones, in press). Misunderstandings between researchers and participants who are culturally different, the paucity of standardized assessments and appropriate treatments for ethnic minority children and adolescents, and the disproportionate extent of abusive or violent behaviors that may occur in impoverished communities may lead to referrals to agencies or reports to authorities that create an unfair burden on poor and culturally diverse children and their families (Jackson, 1993; Scarr, in press; Scott-Jones, in press).

Generalizing the Results of Research to Individuals and Diverse Populations

A cornerstone of experimental methodology involves utilizing large numbers of subjects to generate group means that can then provide information about cognitive, socioemotional, behavioral, or physiological developmental patterns in normative and atypical populations. However, as "knowledge consumers" (e.g., parents, practitioners, and policymakers) come to rely increasingly on the developmental science knowledge base, investigators must examine their own responsibility to limit possible misapplications of findings derived from group data to policy decisions that affect the lives of individual children and youth (see Principle 14, SRCD, 1993).

Overestimating the Relevance of Developmental Research to Practical Applications

The question of how much responsibility scientists should bear for the potential applications of their research has generated heated debate—both about participant autonomy, i.e., whether participants should be informed about the implications of the research to social policy, and about academic freedom, i.e., how much social considerations should influence the pursuit of scientific truth (Fisher & Rosendahl, 1990; Sarason, 1984; Scarr, 1988). It is increasingly clear to action researchers that in order to select appropriate experimental variables in studies of social problems, it is essential to consider the public practices and social circumstances facing members of a target group (Baumrind, 1990; Fisher & Rosendahl, 1990). Less clear is the obligation investigators have to consider the social implications of their research when its relevance to public policy and practice is less direct, that is, what obligation they may have to clarify the limitations of their data or to refrain from proposing action beyond what is known (Hoffman, 1990).

For example, developmental research has demonstrated a relationship between infant-mother attachment and certain social competencies and risks in early childhood (e.g., Bretherton & Waters, 1985; Matas, Arend, & Sroufe, 1978). These important findings are based on comparisons of group means and observations of behavioral patterns in groups of infants identified as securely or insecurely attached, as measured by the Strange Situation procedure (Ainsworth, Blehar, Waters, & Wall, 1978). While developmentalists may recognize the limitations of their research as it applies to an individual child, knowledge consumers (e.g., policymakers, therapists) may not. In fact, policymakers have looked in recent years to attachment research in the infant day-care setting to inform policy decisions concerning maternal leave and government-sponsored child care programs (Belsky, 1986, 1989; Phillips, McCartney, Scarr, & Howes, 1987). Moreover, child and family therapists are increasingly interested in using the Strange Situation procedure as a clinical tool for assessing parent-child relationships.

Since the predictive validity of attachment classifications for later cognitive and socioemotional adjustment in individual infants and infants from different cultural backgrounds is yet to be empirically verified (Bates, Maslin, & Frankel, 1985; Erikson, Sroufe, & Egeland, 1985; Fracasso, Busch-Rossnagel, & Fisher, in press; Grossmann, Grossmann, Huber, & Wartner, 1981; Lamb, 1987; Sagi & Lewkowicz, 1987; Thompson, 1988), it would be inappropriate at this point to assume that knowledge of a specific infant's attachment classification is a valid indicator of individual risk or vulnerability. While researchers cannot be held ethically responsible for the *less obvious* implications others may draw from their findings, they do have an ethical responsibility, in studies of socially sensitive issues (as infant-mother attachment has recently become), to anticipate the potential real world implications of their research and to clarify in their writings and public presentations the extent, as well as the limits, of the application of their findings to practice or social policy (Fisher & Rosendahl, 1990; Hoffman, 1990; Scarr, 1990).

Evaluating Social Programs

Limits of single-strategy interventions. The ethics of applying group findings to the lives of individual children and youth also come into play in empirical evaluations of social programs. Social programs typically prescribe a single intervention to promote the development and welfare of children or adolescents at risk (Baumrind, 1990). The success of the program in meeting its goals is then usually assessed by comparing youth who participated in the program with those who did not. If analyses indicate that the intervention group scored significantly higher on outcome measures than the control or comparison group, the program is judged a success. If the difference between the groups is not statistically significant, the program is considered a failure and may face being abandoned.

While aggregating individual responses to a program increases the power to detect intervention effects, it also decreases the opportunity to learn who is and who is not benefiting from the program. As Lewis points out (Lewis, 1990; Lewis & Sullivan, 1994), basing policy decisions solely on statistical differences between groups runs counter to current empirical information about individual differences and variations in human development. In fact, based upon our knowledge of the contextual factors influencing development, we should assume that individuals who vary in their prior experiences, attitudes, and abilities would in turn vary in their responses to different intervention strategies (Berliner & Cahen, 1978; Lerner & Tubman, 1990). For example, research has indicated that adolescents use drugs for different purposes depending on contextual factors. High school students most at risk for long-term drug problems are school dropouts or adolescents from unmotivated families or families modeling substance abusive behavior (Pentz, 1994). Accordingly, a school-based drug prevention program with an educational/information focus may work well for youth who function well academically and socially, but it may not meet the needs of more troubled and less motivated students.

To assume that single intervention strategies will universally meet the needs of heterogeneous populations can have serious repercussions for children and their families, especially members of disadvantaged or ethnic minority populations (Laosa, 1990). First, adherence to single-strategy interventions runs the risk of leading to the conclusion that the majority of social programs fail (Lewis, 1990). Second, program failure is too often attributed to negative attributes of the target population rather than to limits of the research methodology or the intervention itself. Lastly, a negative program evaluation often results in the decision to abandon the problem, rather than the resolve to find more effective intervention techniques (Scarr, 1990). Thus, evaluations that ignore the interaction of participant variability with program features can result in families most in need being denied services that could promote their children's development.

Multistrategy approaches to program design and evaluation. Given the limitations of single-intervention approaches, developmental scientists must seek to avoid cost-dictated compromises in evaluation research and pursue new program design and evaluation methodology to address the heterogeneity of child and adolescent populations. This can entail three steps: (1) exploring how members of the target population who benefit from the intervention may differ systematically from those who do not; (2) retesting the intervention by evaluating its effect on a new set of participants with the characteristics of those whom step 1 show to benefit from the program; and (3) designing new intervention strategies and outcome measures to fit the characteristics of those who did *not* benefit from the original program (Lewis, 1990; Lewis & Sullivan, 1994).

Such a multistrategy approach does hold the risk that certain subgroups of individuals will be stigmatized or assigned to services that inadvertently keep them out of the mainstream (Laosa, 1990). For example, assigning to drug abuse prevention programs only those high school students who fit a high-risk profile for long-term drug usage can lead to negative peer labeling (Pentz, 1994). One way to avoid such stigmatizing is to structure the intervention to incorporate elective and self-referral procedures (Pentz, 1994).

Abandoning poor intervention strategies, not children's problems. Making the evaluation of social programs more rigorous does not mean that all programs will be evaluated positively. The continued use of ineffective intervention strategies is just as pressing an ethical issue as the abandonment of programs that are helping particular groups within the population. Growing fiscal constraints on federal, state, and community-sponsored social programs call for researchers to evaluate not only the benefits but the human and economic costs of social policies. Money spent on ineffective programs is money lost to effective programs. Therefore, developmentalists have an ethical responsibility to provide program managers and policymakers with fair evaluations of program efficacy. Researchers, and policymakers, should not be afraid to abandon programs that do not work (Scarr, 1990). Rather, investigators can help design multistrategy approaches that will provide policymakers with the information they need to attack, not abandon, the problems facing today's children and youth.

Respecting the Cultural Dimensions of Children's Lives

The principle of respect acknowledges that research participants are themselves moral agents who have the right to judge the purpose and procedures of research in which they are asked to participate (Fisher & Fyrberg, in press). Threats to participant autonomy can stem from two sources: (1) research activities that usurp the parental role, and (2) unrecognized investigator bias.

Intervention goals and participant values. Most developmental intervention programs operate on a two-part assumption: first, that the trajectory of certain behavioral patterns found within certain populations will lead to undesirable outcomes if "uncorrected"; and second, that if "corrected," the new patterns will relate to successful academic and social adaptation (Sigel, 1990). Researchers do not usually question whether the program might be at odds with the family values of participants, nor whether parents or participants consider the "corrected" behaviors or outcomes desirable.

Consider, for example, a social program that aims at increasing the academic assertiveness of Puerto Rican or Korean immigrant children by encouraging the child to make autonomous decisions and question adult decision making. These educational objectives may run counter to the parents' beliefs about how a child should behave, and this in turn can have negative consequences for minor participants. Some evidence suggests that behavioral and emotional problems can indeed emerge in late childhood and adolescence as a result of the breakdown of the hierarchical structure that characterizes many traditional Hispanic and Asian families (Fracasso & Busch-Rossnagel, 1992; McAdoo, 1990; Ryu & Vann, 1992).

It has been suggested that in using developmental knowledge to change individuals, researchers should incorporate both the *etic* (or objective) values of the researcher and the *emic* (or internal) values of the child's culture (Jackson, 1993; Sigel, 1990). This recommendation extends to program evaluations as well; outcome measures that do not take the child's family heritage and values into account may underestimate children's potential and mislead parents, educators, and policymakers.

Investigator biases. Researchers must also appreciate how their own values may be influenced by social attitudes and, in turn, influence policy (Fisher & Tryon, 1990). Throughout the relatively short life span of the science of human development, society's changing attitudes toward children and families have influenced developmental theory, interpretations of data, and the consequent conceptualization of the child and adolescent. The impact of experimenter social values on research and theory is evident, for example, in the central role of religious and racial categories in the pioneering work of G. Stanley Hall (Youniss, 1990) or in the contemporary biases affecting inquiry into the developmental correlates of maternal employment (Hoffman, 1990).

Investigator bias can intrude into every phase of the research enterprise. It can effect (1) constraints on research design (e.g., requiring majority control groups for descriptive research on ethnic minority children's development); (2) the neglect or pursuit of differences between subgroups (e.g., gender or ethnic differences); (3) the entrenchment of particular paradigms (e.g., studying the normative development of majority children and the problematic development of minority children); and (4) the choice of analyses (e.g., whether to pursue an ANOVA three-way interaction). A stereotyping bias can also lead researchers to view ethnic minorities as monolithic, thereby denying members of subgroups their individuality and opportunities for program benefits (Hoffman, 1990; Korchin, 1980; McAdoo, 1990).

Developmentalists must acknowledge their own values and respect those of research participants, if interventions are to be shaped to serve diverse populations. Parents, practitioners, and policymakers will benefit from research that is cautious with respect to investigator values and sensitive to the cultural dimensions of children's lives.

Developmental Psychology and the Courts

Evidence from developmental research is increasingly finding its way into our legal system. Studies relevant to child custody decisions, the rights of minors to consent to medical procedures, and the ability of children to serve as witnesses are just a few examples of the research contribution to the knowledge base that is being incorporated by judges and jurors.

Research designed to address legal questions. Various moral obligations—derived from the principle of beneficence—inevitably emerge as developmentalists seek to generate knowledge to inform the legal system. The impact of applied research on social decisions can be positive or negative (Fisher & Tryon, 1990). Data that are highly reliable and ecologically valid can provide the court with valuable information, but data that show inconsistencies or are generated under experimental conditions vastly different from the real world phenomena of interest can, if utilized by the court, jeopardize the legal rights of defendants or witnesses. Investigators must thus grapple with the difficulty of balancing external validity (ensuring that experimental conditions correspond to the true nature of the phenomena in question) and internal validity (controlling extraneous variables in order to strengthen conclusions to be drawn from findings).

In efforts to achieve this balance, some investigators choose to conduct field studies of children in their real world situations. For example, researchers can assess psychological functioning and family processes before and after custody decisions (Hetherington, Cox, & Cox, 1981; Shaw & Emery, 1987; Wallerstein & Kelly, 1980). Others create well-controlled and forensically realistic contexts to assess the causal role of particular factors in children's behavior. For example, to investigate children's response to leading questions regarding child abuse, Tobey and Goodman (1992) tested in a laboratory setting the accuracy of 4-year-old children's answers to questions about their play with a "baby-sitter" after they had heard a "police officer" accuse her of wrongdoing.

Ecologically valid "staged events" can have ethical costs, however, if debriefing does not disabuse the child of an experimentally induced false belief. Attempting to simulate stressful situations as they naturally occur, while maintaining internal controls at the same time, is often difficult, if not impossible. For example, the researcher cannot take a child away from one of his or her parents to replicate the impact of certain custody decisions; nor would it be ethical to provide adolescents with inaccurate health information to test their decisions about matters critical to their well-being; nor can we subject children to physical or sexual assault to study their reactions to abuse. And while observations in the field can generate rich descriptive information, they cannot provide causal explanations (Goodman & Tobey, 1994). As a consequence, the application of developmental research to legal questions will always be limited.

Given these limitations, developmentalists who conduct empirical studies relevant to the court's interests bear certain ethical responsibilities to assure that (1) the manipulation or setting resembles as closely as possible the phenomena of interest; (2) measures are reliable and experimental conditions adequately controlled; (3) data are presented fairly; and (4) conclusions drawn do not go beyond the data.

The developmental scientist as expert witness. The developmentalist who is called upon to testify to the court faces several ethical challenges. First is the obligation to present research findings honestly and fairly. This can be difficult on two fronts: (1) in that findings are rarely unequivocal, an evenhanded testimony can lead to a mixed conclusion, with seemingly inconsistent data and diverse interpretations—as exemplified by current research on the suggestibility of child witnesses (Ceci & Bruck, 1993a, 1993b); (2) because science and the legal system have different purposes (i.e., searching for truth in the case of science, and securing justice in the case of law [Melton, 1990b; Melton & Limber, 1989]) objective evidence can become distorted in the process. Accordingly, a researcher's first ethical responsibility as an expert witness is to be *competent*—not only in mastery of the developmental literature appropriate to the case at hand, but in familiarity with legal procedures and the skills and knowledge underlying forensic work.

A second challenge facing the developmentalist as expert witness is to avoid the pitfalls of advocating for a particular side of a legal case. Since the expert witness is typically retained by an attorney on one side of a legal debate, it is tempting to view ones duty as providing testimony that will help "win" the case. However, such a position is inappropriate for the expert whose ethical responsibility is rather to be an advocate for the data (Melton, 1990b). This means that the expert is obliged to provide an accurate summary of scientific information to assist the finders of facts (the judge and jurors) in weighing the evidence put before them.

A third ethical responsibility of the expert witness is to avoid generalizing findings derived from group means to the individual case. Unlike other witnesses, the expert witness is allowed to answer "hypothetical" questions. This means that the developmental researcher will be asked to render an opinion about the facts brought before the court. The expert witness must be clear at this point about the difference between the probabilistic nature of science (e.g., that some observed difference between groups of individuals did not arise by chance but rather at some level of probable cause) and the all-or-none nature of justice (e.g., that an individual is or is not found guilty). For example, while one study showed that 84% of the child participants continued to respond accurately about the absence of abuse when asked misleading questions (Tobey & Goodman, 1992), an overgeneralized conclusion based on this finding could be disastrous for a child abuse defendant whose accuser reflects the 13% of those children who do not respond accurately. Similarly, while research has indicated that 50% of 3- and 4-year-olds will occasionally distort the truth to achieve some aim (Ceci & Leichtman, 1992), an attempt to discredit a child's testimony based on this finding can be extremely harmful to a youngster accurately relaying her abusive experiences in court. The developmentalist must clarify for the judge and jury the limitations of applying empirical findings based on probability and group means to the determination of facts about an individual (Melton, 1990b).

Recommendations for Future Policy

Until recently there has been little dialogue or consensus on the complex issues discussed here. However, as developmental research involves more and more high-risk children and adolescents, these issues are becoming more salient. The recommendations that follow are intended to provide a framework for reshaping federal policy on the ethical conduct of research.

Specifying the Conditions of Requiring the Guardian Consent Waiver

Federal guidelines on the waiver of guardian consent must be more specific. At present, exceptions to the general requirements for informed consent put forth by the National Institutes of Health Office for Protection of Research Risks (OPRR) are inconsistently applied across institutional review boards, grant review panels, and the OPRR itself in reviews of research on specific populations (e.g., runaways, children from abusive homes, pregnant adolescents). For example, according to Title 45, Part 46 of the *Code of Federal Regulations* (DHHS, 1991), informed consent may be waived provided the IRB documents that the research investigates factors relevant to public service programs and is subject to government agency approval, or that the research could not be practically carried out without waiver. While this provision appears to give developmental researchers wide latitude in gaining a guardian consent waiver, additional guidelines have been invoked to preclude a majority of such waivers—namely, section 116.d.1 stating that "the research involves no more than *minimal risk* to the subjects" (where the probability and magnitude of harm is not greater than ordinarily encountered in daily life); and section 116.d.2 stating that "the waiver or alteration will not *adversely affect* the rights and welfare of the subjects" (DHHS, 1991, p. 10).

Consistency of reviews. In judging the ethical acceptability of a research proposal, IRBs, grant review panels, and the OPRR often disagree in their evaluations of "minimal risk" or "adverse effect" (Freedman, Fuks, & Weijer, 1993). While the checks and balances provided by the independent evaluation of proposals helps protect participants, it is important that standards not be misapplied. Care must be taken to avoid inadvertently discouraging studies of high-risk children or leading researchers to believe that Title 45, Part 46 of the *Code of Federal Regulations* implies that informed consent is not required, with or without waiver. Here are some issues that need clarification:

- What experimental conditions meet the criteria of "minimal risk"? At present it is not clear whether anonymous surveys or confidential interviews meet federal criteria, nor to what extent the child's age or other child characteristics should influence decisions about minimal risk.

- What types of survey or interview content are to be considered potentially adverse or damaging to minor participants?
- What is the potential negative impact of waiving guardian consent in the case of descriptive versus intervention studies? This is an especially important question for studies of the health-compromising behaviors of adolescent populations.
- How can "risk" be better defined across diverse populations? In that disadvantaged youth may be subject to more everyday risks than their more advantaged counterparts, an "everyday" criterion may inadvertently lead to unjust policies (Kopelman, 1981; Scott-Jones, in press). Furthermore, using an "advantaged" norm to define minimal risk may exclude disadvantaged children from research that could benefit them.
- How is it to be determined that an adolescent does not have a legal guardian, or has a guardian unable to make decisions in the youth's best interest?

Participant advocates. Federal guidelines (DHHS, 1983), recognizing that some children and adolescents do not have a "loving" relationship with a guardian, recommend that under such conditions IRBs appoint an advocate to be present during the informed consent process. Currently, however, the identification and utilization of advocates to represent minor participants has been limited to institutionalized populations.

The critical need for empirical data on the lives of emancipated adolescents without guardians calls for federal support for a system of participant advocacy. This system would require that a researcher who is granted a waiver of guardian consent appoint an individual(s) with no investment in the research project or role in subject recruitment to serve as an advocate for the minors being recruited for participation. The responsibilities of the participant advocate would include evaluating each youth's understanding of informed consent procedures, ensuring that participation is voluntary, checking periodically to determine whether the youth wants to terminate participation, assessing the participant's reactions to research procedures, ensuring that the debriefing addresses all participant questions and concerns, and serving as a liaison to referral services.

Federal Guidelines for Reporting and Referring Research Participants

As discussed earlier, research involving children and adolescents from high-risk environments has the potential to tap sources of psychopathology, psychological distress, or physical vulnerability that jeopardize participants' welfare. Federal guidelines are needed to articulate the investigator's responsibility to share or report such information. These guidelines will help researchers design informed consent procedures and establish, where appropriate, referring and reporting procedures that offer reasonable protections for minor participants and their families.

Clarifying the relationship between federal and state regulations is also critical to the scientific community's ability to design procedures that protect not only the welfare but the privacy rights of participants and their families. Researchers working with high-risk and disadvantaged populations need guidance in the following areas:

- Under what conditions does the federal Certificate of Confidentiality supersede state reporting laws covering current and past abusive behaviors?

- What are the legal risks to researchers who decide to disclose information to protect participant welfare after having obtained a Certificate and having promised participants confidentiality?
- Are researchers under the same obligation as mental health practitioners to report a "suspicion" of child abuse, or does a more stringent criterion for evaluating child jeopardy apply?
- What legal protections, if any, are available to the researcher whose report of a "suspicion" of abuse proves unfounded?

Expanded Federal Funding

In setting child and adolescent welfare as a national priority, the federal government runs the risk of creating false expectations if funding is inadequate to produce ethical practice and empirically sound results (Baumrind, 1990). But federal agencies can help investigators balance their scientific and ethical responsibilities by funding a variety of efforts: (1) innovative ethical practices, (2) multistrategy intervention designs that include basic services to control groups, (3) research on participant reactions to ethical procedures, and (4) applied developmental science training programs.

Innovative practices. Investigators need to build into their projects requests for federal support for design components aimed at protecting the rights and promoting the welfare of research participants. Funding is needed to support more active efforts to obtain guardian consent within the neighborhoods where potential high-risk participants live. This kind of recruitment requires funds for pilot work to establish the most efficacious procedures and to train and pay research personnel for the demanding, labor-intensive task of establishing personal, often door-to-door relationships with the community and with potential participants and their guardians.

Funding is also needed to enhance the potential benefits of participation in research. For example, while subjects typically derive little or no direct benefit from their participation, the debriefing stage of a study offers an ideal opportunity for children and youth to learn how the scientific method is applied to the study of human development, or how to avail themselves of mental health or social services in the community. Supporting budgetary requests for these aspects of research can help introduce a new generation to the methods of human experimentation and link the idea of good science to "doing good."

Multistrategy design. Federal agencies responsible for social service programs should cease expecting and demanding that the first intervention strategy tried be successful, and investigators should strengthen their requests for support of multistrategy intervention designs. Program evaluation of different treatment approaches for children and adolescents varying in cognitive skills, socioemotional adaptation, and cultural and economic background need to be encouraged. More intervention studies requesting and receiving support for offering control group participants *basic* services are also needed.

A "science of scientific ethics." Vulnerable children and youth also stand to benefit from a federally sponsored initiative to investigate participant and guardian responses to current ethical practices and to the enhanced ethical procedures being proposed here (Hoagwood, Jensen, & Leshner, in press). Research is needed to inform the design of ethical practices that will respect participant autonomy, promote participant welfare, and distribute fairly the costs and benefits of scientific inquiry. Following are some of the issues needing investigation:

- What determines guardian consent and participant assent among children and adolescents who differ in actual and perceived risk?
- What is the impact of monetary incentives on participation?
- How do participants and families react to random assignment to control groups in intervention studies, and what characteristics distinguish those who agree and those who refuse to participate in random assignment?
- What type of information do participants and their families want to receive when a study is over?
- Under what conditions do participants and parents believe researchers should report problems identified during the course of research?
- Do members of the focal community and research scientists differ in their evaluation of the scientific value and participant costs of current models of developmental research?

A federal research initiative to generate empirical data on participant perspectives holds the promise not only of building mutual respect between investigator and participant, but of helping to shape ethical practices (Fisher, in press; Fisher & Fyrberg, in press).

Training programs. Expanded federal funding is needed to promote the training of researchers who can work with parents, practitioners, social service agencies, and policy makers (Fisher, et al., 1993; Fisher, Rau, & Colapietro, 1993). For example,

- New curricula should promote ethics education (to include ethical decision making), familiarity with federal and professional ethics codes, and awareness of state and federal laws pertinent to research with high-risk children and youth.
- Students must also acquire the expertise to (1) evaluate the validity of assessment instruments for diverse populations, (2) appraise the multilevel responses of participants to social programs, and (3) evaluate and help develop multistrategy interventions.
- Federal support should encourage curricula that provide students with the skills to disseminate developmental findings clearly and fairly in both legal and political forums.

Conclusion

As developmental scientists leave the laboratory to investigate the real world challenges facing high-risk populations, the ethical obligations to participant welfare become exceedingly complex. These new pressures place the investigator in the "scientist-citizen" dilemma whereby competing obligations to produce scientifically rigorous knowledge and to protect and promote the welfare of research participants must be balanced (Fisher & Tryon, 1990; Scarr, in press; Veatch, 1987). The current cost-benefit approach to ethical decision making often leads investigators to look at scientific rigor and social responsibility as incompatible goals—as either/or possibilities (Fisher, in press). This report presents a framework for integrating these two purposes. Researchers and policymakers must work together to develop procedures and experimental designs that enable investigators to fulfill their obligation to a research ethic of scientific responsibility and care.

References

Achenbach, T. M., & Edelbrock, C. (1983). *Manual for the Child Behavior Checklist and Revised Child Behavior Profile*. Burlington: University of Vermont, Department of Psychiatry

Ainsworth, M. D., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment.* Hillsdale, NJ: Erlbaum.

American Educational Research Association. (1992). *Ethical standards of the American Educational Research Association*. Washington, DC : Author.

American Psychological Association. (1992). Ethical principles of psychologists and code of conduct. *American Psychologist, 47,* 1597-1611.

Appelbaum, P. S., & Rosenbaum, A. (1989). *Tarasoff* and the researcher: Does the duty to protect apply in the research setting? *American Psychologist, 44,* 885–894.

Asher, S. R. (1993, March). *Inviting children to self-refer.* Paper presented as part of the invited symposium on "Ethical issues in the reporting and referring of research participants" (C. B. Fisher, chair) at the biennial meeting of the Society for Research in Child Development, New Orleans.

Asher, S. R., & Dodge, K. A. (1986). Identifying children who are rejected by their peers. *Developmental Psychology, 22,* 444–449.

Bates, J. E., Maslin, C. A., & Frankel, K. A. (1985). Attachment, security, and temperament as predictors of behavior: Problem ratings at three years. In I. Bretherton & E. Waters (Eds.), *Growing points in attachment theory* (pp. 167–193). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Baumrind, D. (1990). Doing good well. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 17–28). Norwood, NJ: Ablex.

Beauchamp, T. L., & Childress, J. F. (1979). *Principles of biomedical research.* New York: Oxford University Press.

Belsky, J. (1986). Infant day care: A cause for concern? *Zero to Three, 6,* 1–9.

Belsky, J. (1989). Infant-parent attachment and day care: In defense of the strange situation. In J. Lande, S. Scarr, & N. Gunzenbauser (Eds.), *Caring for children: Challenge to America* (pp. 23–48). Hillsdale, NJ: Erlbaum.

Berliner, D., & Cahen, L. S. (1978). Trait-treatment interactions and learning. In F. N. Kerlinger (Ed.), *Review of research in education* (Vol. 1). Ithaca, NY: Peacock.

Bretherton, I., & Waters, E. (Eds). (1985). Growing points in attachment theory. *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Brooks-Gunn, J., & Rotheram-Borus, M. J. (in press). Adolescent girls and mothers: Rights and privacy in research. *Ethics & Behavior.*

Carroll, M. A., Schneider, H. G., & Wesley, G. R. (1985). *Ethics in the practice of psychology*. Englewood, NJ: Prentice-Hall.

Ceci, S. J., & Bruck, M. (1993a). Suggestibility of the child witness: A historical review and synthesis. *Psychological Bulletin, 113,* 403–409.

Ceci, S. J., & Bruck, M. (1993b). Child witnesses: Translating research into policy. *Social Policy Report, 7,* 1-31.

Ceci, S. J., & Leichtman, M. D. (1992). "I know that you know that I know that you broke the toy": A brief report of recursive awareness among 3-year-olds. In S. J. Ceci, M. D. Leichtman, & M. E. Putnick (Eds.), *Cognitive and social factors in early deception.* Hillsdale, NJ: Erlbaum.

Chase-Lansdale, P. L., & Brooks-Gunn, J. (1994). Correlates of adolescent pregnancy and parenthood. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 207–236). New York: McGraw-Hill.

Child Abuse Prevention and Treatment Act, 42 U.S.C. Article 5101 (1976 & Supp. 1987).

Damon, W. (1983). *Social and personality development: Infancy through adolescence.* New York: Norton.

Department of Health and Human Services. (1983, March 8). Additional protections for children involved as subjects in research. *Federal Register, 48*(46), 9814–9820.

Department of Health and Human Services. (1991, August). Title 45 Public Welfare, Part 46, *Code of Federal Regulations, Protection of Human Subjects.*

Department of Health, Education, and Welfare. (1978). *The Belmont Report: Ethical principles and guidelines* *for the protection of human subjects of research* (DHEW Publication No.[OS] 78–0012, Appendix I, DHEW Publication No. [OS] 78–0013, Appendix II, DHEW Publication [OS] 78–0014). Washington, DC: U.S. Government Printing Office.

Edelbrock, C. (1994). Assessment of child psychopathology. In C. B. Fisher & R. M. Lerner (Eds.). *Applied developmental psychology* (pp. 294–314). New York: McGraw-Hill.

Erickson, M. F., Sroufe, A. L., & Egeland, B. (1985). The relationship between the quality of attachment and behavior problems in preschool in a high-risk sample. In I. Bretherton & E. Waters (Eds.), *Growing points of attachment theory and research.* (pp. 147–186). *Monographs of the Society for Research in Child Development, 50*(1–2, Serial No. 209).

Faden, R. R., & Beauchamp, T. L. (1986). *A history and theory of informed consent.* New York: Oxford University Press.

Fisher, C. B. (1991). Ethical considerations for research on psychosocial interventions for high-risk infants and children. *Register Report, 17,* 9-12.

Fisher, C. B. (1993). Joining science and application: Ethical challenges for researchers and practitioners. *Professional Psychology: Research and Practice, 24,* 378–381.

Fisher, C. B. (in press). Reporting and referring research participants: Ethical challenges for investigators studying children and youth. *Ethics & Behavior.*

Fisher, C. B., & Brennan, M. (1992). Application and ethics in developmental psychology. In D. L. Featherman, R. M. Lerner, & M. Perlmutter (Eds.), *Life-span development and behavior* (pp. 189–219). Hillsdale, NJ: Erlbaum.

Fisher, C. B., & Fyrberg, D. (in press). Participant partners: College students weigh the costs and benefits of deceptive research. *American Psychologist.*

Fisher, C. B., & Lerner, R. M. (1994). Foundations of applied developmental psychology. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 3–20). New York: McGraw-Hill.

Fisher, C. B., Murray, J. P., Dill, J. R., Hagen, J. W., Hogan, M. J., Lerner, R. M., Rebok, G. W., Sigel, I. E., Sostek, A. M., Smyer, M. A., Spencer, M. B., & Wilcox, B. (1993). The national conference on graduate education in the applications of developmental science across the lifespan. *Journal of Applied Developmental Psychology, 14,* 1–10.

Fisher, C. B., Rau, J. B., & Colapietro, E. (1993). The Fordham University doctoral specialization in applied developmental psychology. *Journal of Applied Developmental Psychology, 14,* 289–302.

Fisher, C. B., & Rosendahl, S. A. (1990). Psychological risks and remedies of research participation. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 43–60). Norwood, NJ: Ablex.

Fisher, C. B., & Tryon, W. W. (1988). Ethical issues in the research and practice of applied developmental psychology. *Journal of Applied Developmental Psychology, 9,* 27–39.

Fisher, C. B., & Tryon, W. W. (1990). Emerging ethical issues in an emerging field. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 1–14). Norwood, NJ: Ablex.

Fracasso, M. P., & Busch-Rossnagel, N. A. (1992). Parents and children of Hispanic origin. In M. E. Procidano & C. B. Fisher (Eds.), *Contemporary families: A handbook for school professionals* (pp. 83–98). New York: Teachers College Press.

Fracasso, M. P., Busch-Rossnagel, N. A., & Fisher, C. B. (in press). The relationship of maternal behaviors and acculturation to the quality of infant attachment in Puerto Rican and Dominican infants. *Hispanic Journal of Behavioral Science.*

Freedman, B. (1975). A moral theory of informed consent. *Hastings Center Report, 5,* 32–39.

Freedman, B., Fuks, A., & Weijer, C. (1993). *In loco parentis*: Minimal risk as an ethical threshold for research upon children. *Hastings Center Report, 23*(2), 13–19.

Gaylin, W., & Macklin, R. (1982). *Who speaks for the child: The problems of proxy consent.* New York: Plenum.

Gibbs, J. T. (1990). Mental health issues of black adolescents: Implications for policy and practice. In A. R. Stiffman & L. E. Davis, (Eds.), *Ethnic issues in adolescent mental health* (pp. 21–52). Newbury Park, CA: Sage Publications.

Goodman, G. S., & Tobey, A. E. (1994). Memory development within the context of child sexual abuse investigations. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 46–75). New York: McGraw-Hill.

Graham, S. (1992). "Most of the subjects were white and middle class": Trends in published research on African Americans in selected APA journals, 1970–1989. *American Psychologist, 47,* 629–639.

Grossmann, K. E., Grossmann, K., Huber, F., & Wartner, U. (1981). German children's behavior towards their mothers at 12 months and their fathers at 18 months in Ainsworth's Strange Situation. *International Journal of Behavioral Development, 4,* 157–181.

Hagen, J. W., Paul, B., Gibb, S., & Wolters, C. (1990, March). *Trends in research as reflected by publications in child development: 1930–1989.* Paper presented at the biennial meeting of the Society for Research in Adolescence, Atlanta, GA.

Hendren, R. L. (1991). Determining the need for inpatient treatment. In R. L. Hendren & I. N. Berlin (Eds.), *Psychiatric inpatient care of children and adolescents: A multi-cultural approach* (pp. 37–64). New York: Wiley.

Hetherington, E. M., Cox, M., & Cox, R. (1981). Effects of divorce on parents and children. In M. E. Lamb (Ed.), *Nontraditional families* (pp. 35–58). Hillsdale, NJ: Erlbaum.

Hoagwood, K. (in press). The Certificate of Confidentiality at NIMH: Applications and implications for service research with children. *Ethics & Behavior.*

Hoagwood, K., Jensen, P., & Leshner, A., (in press). Ethical issues in research on child and adolescent mental disorders: Implications for a science of scientific ethics. *Innovative Services for Difficult to Treat Populations.* Washington, DC: American Psychiatric Press.

Hoffman, L. W. (1990). Bias and social responsibility in the study of maternal employment. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 253–272). Norwood, NJ: Ablex.

Holder, A. R. (1981). Can teenagers participate in research without parental consent? *IRB: Review of Human Subjects Research, 3*(2), 5–7.

Horowitz, F. D., & O'Brien, M. (1989). In the interest of the nation: A reflective essay on the state of our knowledge and challenges before us. *American Psychologist, 44,* 441–445.

Jackson, J. F. (1993). Multiple caregiving among African Americans and infant attachment: The need for an emic approach. *Human Development, 36,* 87–102.

Kandel, D., Davies, M., Karus, D., & Yamaguchi, K. (1986). The consequences in young adulthood of adolescent drug involvement. *Archives of General Psychiatry, 43,* 746–754.

Keith-Spiegel, P. C. (1983). Children and consent to participate in research. In G. P. Melton, G. P. Koocher, & M. J. Saks (Eds.), *Children's competence to consent* (pp.179–211). New York: Plenum.

Koocher, G. P., & Keith-Spiegel, P. C. (1990). *Children, ethics, and the law.* Lincoln: University of Nebraska Press.

Kopelman, L. (1981). Estimating risk in human research. *Clinical Research, 29,* 1–8.

Korchin, S. (1980). Clinical psychology and minority problems. *American Psychologist, 35,* 890–901.

Kovacs, M. (1981). Rating scales to assess depression in school children. *Acta Paedopsychiatrica, 46,* 305–315.

Lamb, M. E. (1987). Predictive implications of individual differences in attachment. *Journal of Consulting and Clinical Psychology, 55,* 817–824.

Laosa, L. (1990). Population generalizability, cultural sensitivity, and ethical dilemmas. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 227–252). Norwood, NJ: Ablex.

Lerner, R. M., & Tubman, J. G. (1990). Plasticity in development: Ethical implications for developmental interventions. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 113–132). Norwood, NJ: Ablex.

Levine, R. (1986). *Ethics and regulation of clinical research* (2nd edit.). Baltimore-Munich: Urban & Schwarzenberg.

Lewis, M. (1990). Assessing social intervention: Scientific and social implications. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 81–92). Norwood, NJ: Ablex.

Lewis, M., & Sullivan, M. W. (1994). Developmental intervention in the lives of infants and parents. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 375–406). New York: McGraw-Hill.

Liss, M. (in press). State and federal laws governing reporting for researchers. *Ethics & Behavior.*

Macklin, R. (1981). "Due" and "undue" inducements: On paying money to research subjects. *IRB: Review of Human Subjects Research, 3*(5), 1–6.

Matas, L., Arend, R., & Sroufe, A. L. (1978). Continuity of adaptation in the second year: The relationship between quality of attachment and later competence. *Child Development, 49,* 547–556.

McAdoo, H. P. (1990). The ethics of research and intervention with ethnic minority parents and their children. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 273–284). Norwood, NJ: Ablex.

McLoyd, V. C., & Randolph, S. M. (1984). The conduct and publication of research on Afro-American children: A content analysis. *Human Development, 27,* 65–75.

Melton, G. B. (1990a). Certificates of Confidentiality under the Public Health Services Act: Strong protection but not enough. *Violence and Victims, 5*, 67–71.

Melton, G. B. (1990b). Ethical dilemmas in playing by the rules: Applied developmental research and the law. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 145–162). Norwood, NJ: Ablex.

Melton, G. B., Koocher, G. P., & Saks, M. J. (1983). *Children's competence to consent*. New York: Plenum.

Melton, G. B., & Limber, S. (1989). Psychologists' involvement in cases of child maltreatment: Limits of role and expertise. *American Psychologist, 44*, 1225–1233.

Nuremberg Code. (1946). *Journal of the American Medical Association, 132*, 1090.

Pentz, M. A. (1994). Primary prevention of adolescent drug abuse. In C. B. Fisher & R. M. Lerner (Eds.), *Applied developmental psychology* (pp. 435–474). New York: McGraw-Hill.

Phillips, D., McCartney, K., Scarr, S., & Howes, C. (1987). Selective review of infant day care research: A cause for concern? *Zero to Three, 7*, 18–21.

Psychological Corporation. (1991). *Wechsler Intelligence Scale for Children–R*. Houston, TX: Author.

Rawls, J. (1971). *A theory of justice*. Cambridge: Harvard University Press.

Ryu, J. P., & Vann, B. H. (1992). Korean families in America. In M. E. Procidano & C. B. Fisher (Eds.), *Contemporary families: A handbook for school professionals* (pp. 117–134). New York: Teachers College Press.

Sagi, A., & Lewkowicz, K. S. (1987). A cross-cultural evaluation of attachment research. In L. W. C. Tavecchio & M. H. van Ijzendoorn (Eds.), *Attachment in social networks* (pp. 427–455). Amsterdam: North-Holland.

Sarason, S. (1984). If it can be studied or developed, should it? *American Psychologist, 39*, 739–754.

Scarr, S. (1988). Race and gender as psychological variables: Social and ethical issues. *American Psychologist, 43*, 56–59.

Scarr, S. (1990). Ethical dilemmas in recent research: A personal saga. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 29–42). Norwood, NJ: Ablex.

Scarr, S. (in press). Ethical problems in research on risky behaviors and risky populations. *Ethics & Behavior*.

Scott-Jones, D. (in press). Ethical issues in reporting and referring in research with minority and low-income populations. *Ethics & Behavior*.

Shaw, D. S., & Emery, R. E. (1987). Parental conflict and other correlates of the adjustment of school-age children whose parents have separated. *Journal of Abnormal Child Psychology, 15*, 269–281.

Sigel, I. E. (1990). Ethical concerns for the use of research findings in applied settings. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 133–142). Norwood, NJ: Ablex.

Society for Research in Child Development. (1993). Ethical standards for research with children. In *Directory of Members* (pp. 337-339). Ann Arbor, MI: Author.

Tarasoff v. *Regents of the University of California*. (1976). 131 Cal. Rptr.14, 551, P.2nd 553.

Thompson, R. A. (1988). The effects of infant day care through the prism of attachment theory: A critical appraisal. *Early Childhood Research Quarterly, 3*, 273–282.

Thompson, R. A. (1990). Vulnerability in research: A developmental perspective on research risk. *Child Development, 61*, 1–16.

Tobey, A. E., & Goodman, G. S. (1992). Children's eyewitness memory: Effects of participation and forensic context. *Child Abuse & Neglect, 16*, 779–796.

Veatch, R. M (1981). Commentary: Beyond consent to treatment. *IRB: A Review of Human Subjects Research, 3*(5), 7–8.

Veatch, R. M. (1987). *The patient as partner*. Bloomington: Indiana University Press.

Wallerstein, J. S., & Kelly, J. B. (1980). *Surviving the breakup*. New York: Basic Books.

Wasik, B. H., Ramey, C. T., Bryant, D. M., & Sparling, J. J. (1990). A longitudinal study of two early intervention strategies: Project CARE. *Child Development, 61*, 1682–1696.

Youniss, J. (1990). Cultural forces leading to scientific developmental psychology. In C. B. Fisher & W. W. Tryon (Eds.), *Ethics in applied developmental psychology: Emerging issues in an emerging field* (pp. 285–300). Norwood, NJ: Ablex.

About the Author

Celia B. Fisher, Ph.D., is professor and director of the Graduate Program in Developmental Psychology at Fordham University. She is a member of the Ethics Committee of the American Psychological Association. Dr. Fisher has also served as chair of the New York State Board for Psychology and as chair of the Committee for Ethical Conduct in Child Development Research for the Society for Research in Child Development.

Recent Issues

Volume VI (1992)

No. 1. (Spring) *Mother-headed families: An international perspective and the case of Australia.*
Ailsa Burns.

No. 2. (Summer) *Testing in American Schools: Issues for research and policy.*
Patricia Morison.

No. 3. (Fall) *The states and the poor: Child poverty rises as the safety net shrinks.*
Julie Strawn.

No. 4. (Winter) *Crack's children: The consequences of maternal cocaine abuse.*
Theresa Lawton Hawley and Elizabeth Disney.

Volume VII (1993)

No. 1. *Canadian special education policies: Children with learning disabilities in a bilingual and multicultural society.*
Linda S. Siegel and Judith Wiener.

No. 2. *Using research and theory to justify and inform Head Start expansion.*
Edward Zigler and Sally J. Styfco

No. 3. *Child witnesses: Translating research into policy.*
Stephen J. Ceci and Maggie Bruck.

Subscriptions ($10 to nonmembers of SRCD) and single issues ($3.50) are available from:

SRCD Executive Office • University of Michigan
300 North Ingalls, 10th floor • Ann Arbor, MI 48109-0406

Social Policy Report *is a quarterly publication of the Committee on Child Development, Public Policy and Public Information of the Society for Research in Child Development.*

Copyright of the articles published in the Report *is maintained by the originator[s] of the work. Statements appearing in the* Report *are the views of the author and do not imply endorsement by the Editor or by SRCD.*

Editor: Nancy G. Thomas
Phone: (303) 925-5516
Fax: (303) 925-9570

SRCD Executive Office
University of Michigan
300 N. Ingalls 10th Floor
Ann Arbor, MI 48109-0406

SOCIAL POLICY REPORT